Logistic-ELM: a novel fault diagnosis method for rolling bearings

Zhenhua Tan · Jingyu Ning · Kai Peng · Zhenche Xia · Danke Wu

Received: 25 April 2022 / Accepted: 30 September 2022 / Published online: 25 October 2022
© The Author(s), under exclusive licence to The Brazilian Society of Mechanical Sciences and Engineering 2022

Abstract
The fault diagnosis of rolling bearings is a critical technique to realize predictive maintenance for mechanical condition monitoring. In real industrial systems, the main challenges for the fault diagnosis of rolling bearings pertain to the accuracy and real-time requirements. Most existing methods focus on ensuring the accuracy, and the real-time requirement is often neglected. In this paper, considering both requirements, we propose a novel fast fault diagnosis method for rolling bearings, based on extreme learning machine (ELM) and logistic mapping, named logistic-ELM. First, we identify 14 kinds of time-domain features from the original vibration signals according to mechanical vibration principles and adopt the sequential forward selection strategy to select optimal features from them to ensure the basic predictive accuracy and efficiency. Next, we propose the logistic-ELM for fast fault classification, where the biases in ELM are omitted and the random input weights are replaced by the chaotic logistic mapping sequence which involves a higher uncorrelation to obtain more accurate results with fewer hidden neurons. We conduct extensive experiments on the rolling bearing vibration signal dataset of the Case Western Reserve University bearing data centre. The experimental results show that the proposed approach outperforms existing state-of-the-art comparison methods in terms of the predictive accuracy, and the highest accuracies are 100%, 99.71%, 98%, 100%, 100%, and 100%, respectively, in seven separate sub data environments. Moreover, in terms of the runtime cost, the experimental results indicate that the proposed logistic-ELM can predict the fault in 40 ms with a high accuracy, up to 21-1858 times more rapid than existing methods based on support vector machine, convolutional neural network and multi-scale entropy. Other experiments of fault diagnosis of the rolling bearings under four different loads also indicate that the logistic-ELM can adapt to different operation conditions with high efficiency. The relevant code is publicly available at https://github.com/TAN-OpenLab/logistic-ELM.

Keywords Fault diagnosis · Rolling bearing · Extreme learning machine · Logistic mapping · Machine learning

1 Introduction
Rolling bearings are common components of rotating machinery, which is widely used in the electric power, metallurgical, petrochemical, machinery manufacturing, aerospace domain and other fields. Due to the long-term operation of rolling bearings in harsh operating environments, the probability of fault occurrence is extremely high. For example, in a wind turbine gearbox, bearing failure accounts for 76% of all failures [1]. Rolling bearing faults often trigger a series of chain reactions, ranging from the interruption of the production process, which leads to economic loss, to catastrophic accidents, which may lead to casualties. Therefore, the fault diagnosis of rolling bearings is of significance to both the industry and academia.
1.1 Existing situation

From the 1960s, researchers dedicated to develop fault diagnosis of rolling bearings through different mechanisms such as vibration, static electricity, temperature and ferrographs. Initially, people diagnosed rolling bearing faults by developing simple instruments, such as [2–5], which were usually based on some specific physical principles but with lower accuracy. With the rapid development of the signal analysis theory, especially after the emergence of the fast Fourier transform [6], people adopted a variety of spectrum analysers to diagnose bearing faults. Schemes based on feature extraction from time domain or frequency domain improved the performance of fault diagnosis. Time-domain feature analysis methods usually determine the possible running states of rolling bearings by calculating and analysing parameters and indexes of the vibration signals with various time-domain feature parameters, such as [7–13], while frequency-domain feature analysis methods focus on separating or strengthening the frequency components of the fault signals, such as [14–20], generally with higher accuracies than those of time-domain-based methods. The time-frequency analysis methods combine the two to form a joint function, which could describe the nonlinear and non-stationary dynamic signals of complex mechanical equipment, such as SVM-based methods [21–26], Bayesian classifier [27], and neural network-based algorithms [28–35]. More details are described in Sect. 2.

1.2 Motivation

In real industrial systems, the main challenges for the fault diagnosis of rolling bearings pertain to the accuracy and real-time requirements. However, most existing methods focus on ensuring the accuracy, and the real-time requirement is often neglected. Traditional research based on time or frequency domain usually focused on examining the properties of a certain indicator and returned an approximate diagnosis but not real-time classification of diagnosis. Machine learning- or deep learning-based methods usually directly input the vibration signals into a classifier but did not consider relevance of and redundancy in the feature indicators, owing to which the classification may need more time cost before the diagnosis output. However, in practical applications, fault diagnosis is accompanied through mechanical work where the response speed must be sufficiently high. In order to reduce the computation cost, multiple ELM-based diagnosis methods have been proposed, whereas the randomness of the input weights generated for ELM caused inherent uncertainty for the diagnosis. Consequently, we need a more accurate and faster classification for fault diagnosis of rolling bearings, which takes advantages of existing technologies as well as eliminates the uncertainty of ELM.

1.3 Solution and contribution

To improve the above-mentioned, this paper proposes a novel fault diagnosis method for rolling bearings, known as the logistic-ELM, a variant of ELM. To obtain the information contained in the vibration signals as fully as possible and ensure the accuracy, we first model the original vibration signals and extract a group of key features in the field of mechanical vibration based on the principle of mechanical signals, then utilize SFS to reduce the information redundancy may exist among the features and select the most valuable features for the fault diagnosis. To obtain a more stable predictive accuracy, we use pseudo-random sequence generated by a chaotic logistic mapping [36] to replace the random input weight matrix in ELM. Using the combination of the logistic mapping and the ELM algorithm, a pseudo-random sequence is generated through logistic mapping, and the values in the sequence are arranged into a matrix as the input weight matrix of the ELM. And the bias matrix of the ELM is set as zero. Under the same diagnosis environment, the input weight matrix needs to be generated only once, and the input weights can be reused in the subsequent diagnosis. The main contributions of this paper are as follows:

1. We consider 14 kinds of physical mechanical principles to extract time-domain features from vibration signals, which can be calculated fast during fault diagnosis. Multiple signal features are jointly considered in the processes of fault diagnosis, and the advantages of each feature are cross used to ensure the predictive accuracy and efficiency.
2. The proposed logistic-ELM fully exploits the internal uncorrelation of the chaotic logistic mapping, and enhances the difference between the hidden neurons by generating input weight matrix through logistic mapping, further ensuring the runtime efficiency and predictive accuracy in fault diagnosis of rolling bearings.
3. Experimental results show that the proposed logistic-ELM outperforms existing state-of-the-art comparison methods in terms of the predictive accuracy and can realize the fault diagnosis with a reasonable predictive accuracy in 40 ms, up to 21-1858 times more rapid than existing comparisons.

The remaining article is organized as follows. The related work and basic principle of the ELM are briefly described in Sect. 2. The design of the proposed method is described in Sect. 3. We use the rolling bearing vibration signal data set of the CWRU to perform a simulation to verify and analyse
the proposed method, as described in Sect. 4. The concluding remarks are presented in Sect. 5.

2 Related work and preliminary

2.1 Related work

Research on the fault diagnosis of rolling bearings began in the 1960s [2, 3]. Initially, people diagnosed bearing faults by using simple instruments such as listening sticks. In 1971, a Swedish instrument company first developed the shock pulse measurement (SPM) technique to detect bearing faults [4]. The SPM is based on the principle that the bearing failure impact can induce the resonance of the system, and bearing faults are diagnosed according to the magnitude of the impact pulse. However, this technique can only identify whether the bearing is faulty or not but the fault types cannot be classified, and the technique exhibits a weak anti-interference ability and low stability. In general, the identification of a fault occurrence in a rolling bearing and classification of the fault type can be realized considering the feature frequency of the faults in vibration signal spectra. In 1974, [5] presented a patent for a resonance demodulation analysis system. This system processes the fault signal considering the resonance and by using a band-pass filter, which enhances the signal-to-noise ratio of the signal. Moreover, the accuracy of the diagnosis result is increased through envelope processing. However, a certain experience is necessary to determine the centre frequency and bandwidth of the filter, and the efficiency of diagnosis is low. In recent years, with the rapid advancement of computer technologies, various new signal processing methods have emerged. Many scholars have successively proposed computer-centric rolling bearing fault diagnosis methods, and a fault diagnosis scheme based on feature extraction and fault classification has been established.

Feature extraction approaches can be primarily divided into three categories, namely time-domain, frequency-domain, and time-frequency-domain feature extraction. The time-domain feature extraction is based on various time-domain feature parameters and performance indexes of the vibration signals. By analysing the parameters and indexes, the fault of the equipment can be effectively determined in an intuitive and understandable manner [7]. In general, traditional time-domain features can be divided into dimensional and non-dimensional parameter indexes, which are easy to calculate and can reflect the working state of the bearings to a certain extent. The traditional research is focused on examining the properties of a single parameter, and fault diagnosis schemes combining multiple traditional time-domain features remain to be developed. In addition to traditional time-domain feature extraction methods, the information entropy theory has been applied for time-domain feature extraction [8–13, 37]. However, the computation time of the information entropy algorithm is usually large because of its complexity.

In contrast to the time-domain analysis, frequency-domain analysis is focused on separating or strengthening the frequency components of the fault signal. This analysis technique is widely adopted in mechanical fault diagnosis because the distribution of the frequency components of the signal in the spectrum can be obtained through the fast Fourier transform, which is more intuitive than the time-domain waveform. The Fourier transform performs statistical averaging on the signal in the time domain by integral transformation, thereby smoothening the non-stationary components in the signal. Although this approach can reflect the frequency information in the signal, it involves certain limitations as the change in the signal over time is not clarified. At present, the commonly used frequency-domain feature extraction methods include power spectrum analysis, envelope spectrum analysis [14, 15], cepstrum analysis [16, 17], high-order spectrum analysis [18] and spectral kurtosis analysis [19, 20].

In the time-frequency analysis, the time and frequency domains are combined to form a joint function to describe the nonlinear and non-stationary dynamic signals of complex mechanical equipment. Common time-frequency analysis methods include the Wigner Ville distribution (WVD), short time Fourier transform (STFT), and wavelet transform (WT). Many scholars proposed adaptive decomposition methods by studying the laws and features of complex signals. In 1998, [38] proposed a method to analyse a non-stationary signal, named empirical mode decomposition (EMD). Moreover, the authors proposed a popular time-frequency analysis method known as the Hilbert-Huang transform, based on the Hilbert transform. In 2005, [39] proposed the local mean decomposition method (LMD). The LMD can decompose a complex multi-component signal into several product functions (PF components) and a residual component. Each PF component consists of a product of an envelope signal and a pure FM signal. Compared with the EMD, the LMD method involves prompt convergence. Moreover, the latter approach can effectively reduce the errors generated in the iteration process and enhance the decomposition accuracy.

Fault classification are usually based on machine learning, and three types of algorithms are widely used. The first type pertains to the SVM [21]. For example, [22] proposed a fault diagnosis algorithm in 2011, in which the feature vectors of the vibration signals are extracted through the wavelet transform, and fault classification is conducted using the SVM algorithm. The second type pertains to the Bayesian classifier [27], which is easy to understand and exhibits a high learning efficiency. However, the algorithm is implemented assuming the independence and normality of the
independent and continuous variables, respectively, which may affect the accuracy of the algorithm to a certain extent. The third type pertains to neural-network-based algorithms, including the back propagation neural network [40], CNN [28, 41], and ELM [29]. For example, the fault diagnosis algorithm proposed by [30] in 2020 involved 2000 input neurons and a hidden layer containing 10 neurons. Reference [31] calculated the time-frequency features of vibration signals by realizing the continuous wavelet transform (CWT) of the complex Morlet wavelet and performing jointed time-frequency analysis (JTFA). Next, the CNN input was obtained through normalization, and the CNN was trained using a time-frequency map with labels. Finally, the trained model was used for fault diagnosis. The HGSA-ELM algorithm [42] proposed by Luo M et al. in 2016 optimized the input weights and bias of the ELM through the real-valued gravitational search algorithm (RGSA), and the binary-valued of the GSA (BGSA) was used to select the valuable features from a compound feature set. Three types of fault features, specifically, time and frequency features, energy features and singular value features, were extracted to compose the compound feature set by applying ensemble empirical mode decomposition (EEMD).

2.2 Preliminary: ELM

The ELM [29] algorithm can be used for a single hidden layer feedforward neural network (SLFN), as shown in Fig. 1.

Traditional feedforward neural networks involve notable limitations, such as a low speed, presence of local minimum, improper learning rate and overfitting. The ELM randomly generates the connection weight between the input and hidden layers, and this weight does not need to be adjusted during the training process. Only the number of hidden layer neurons must be set to obtain the unique optimal solution. Compared with the previous traditional training methods, the ELM exhibits a higher learning speed and generalization performance.

For \( N \) arbitrary distinct samples \((x_i, t_i)\), where \( x_i = [x_{i1}, x_{i2}, \ldots, x_{im}]^T \in \mathbb{R}^m \) and \( t_i = [t_{i1}, t_{i2}, \ldots, t_{im}]^T \in \mathbb{R}^m \), standard SLFNs with \( \tilde{N} \) hidden neurons and activation function \( g(x) \) can be mathematically modelled as

\[
\sum_{i=1}^{N} \beta_i g(w_i \cdot x_j + b_i) = o_j, j = 1, \ldots, N
\]  

(1)

where \( w_i = [w_{i,1}, w_{i,2}, \ldots, w_{i,m}]^T \) is the weight vector connecting the \( i \)th hidden neuron and input neurons, and \( b_i \) is the threshold of the \( i \)th hidden neuron. \( w_i \cdot x_j \) denotes the inner product of \( w_i \) and \( x_j \).

The standard SLFNs with \( \tilde{N} \) hidden neurons and activation function \( g(x) \) can approximate the \( N \) samples with zero error, specifically, \( \sum_{i=1}^{N} \| o_j - t_j \| = 0 \). Consequently, there exist \( \beta_i, w_i \) and \( b_i \) such that

\[
\sum_{i=1}^{\tilde{N}} \beta_i g(w_i \cdot x_j + b_i) = t_j, j = 1, \ldots, N
\]  

(2)

\( N \) can be compactly expressed as

\[
H \beta = T
\]  

(3)

where

\[
H = \begin{bmatrix} 
  g(w_1 \cdot x_1 + b_1) & \cdots & g(w_{\tilde{N}} \cdot x_1 + b_{\tilde{N}}) \\
  \vdots & \ddots & \vdots \\
  g(w_1 \cdot x_N + b_1) & \cdots & g(w_{\tilde{N}} \cdot x_N + b_{\tilde{N}}) 
\end{bmatrix}_{N \times \tilde{N}}
\]  

(4)

\[
\beta = \begin{bmatrix} 
  \beta_1^T \\
  \vdots \\
  \beta_{\tilde{N}}^T 
\end{bmatrix}_{\tilde{N} \times m}, \quad T = \begin{bmatrix} 
  t_1^T \\
  \vdots \\
  t_N^T 
\end{bmatrix}_{N \times m}
\]  

(5)

Identifying the specific \( \hat{\beta}_i, \hat{\beta}_L \) and \( \hat{\beta}_i (i = 1, \ldots, \tilde{N}) \) is equivalent to minimizing the cost function

\[
E = \sum_{j=1}^{N} \left( \sum_{i=1}^{\tilde{N}} \hat{\beta}_i g(w_i \cdot x_j + b_i) - t_j \right)^2
\]  

(6)

The input weights and hidden layer biases of the SLFNs do not need be adjusted and can be arbitrarily assigned. For fixed input weights \( w_i \) and hidden layer biases \( b_i \), training an SLFN is equivalent to finding a least-squares solution \( \hat{\beta}_i \) of

Fig. 1 Framework of the ELM
the linear system $H\beta = T$. The smallest norm least-squares solution of the above linear system is

$$\hat{\beta} = H^+ T$$

(7)

where $H^+$ is the Moore-Penrose generalized inverse of matrix $H$.

### 3 Proposed method

In this section, we describe the design of proposed logistic-ELM. Let $D : \{x_i, y_i\}_{i=1}^N$ indicate distinct samples, where $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^T \in \mathbb{R}^n$ is a n-dimension vibration signal, with corresponding fault label $y_i \in \{1, \ldots, m\}$. For simplicity, let $X = [x_1, \ldots, x_N]^T \in \mathbb{R}^{N \times n}$ represent vibration signals, and $Y = [y_1, \ldots, y_N]^T \in \mathbb{R}^N$ denote the fault labels of $X$. Let $c_i \in 1, \ldots, m$ be the classification result of $x_i$, which is the factual output of the classifier. $C = [c_1, \ldots, c_N]^T \in \mathbb{R}^N$ represents the classification result of $X$. Let $F = [f_1, \ldots, f_N]^T \in \mathbb{R}^{N \times k}$ be the feature matrix of $X$, where $f_i$ is the feature vector of $x_i$. The framework of the proposed method is shown in Fig. 2. Based on the original vibration signals $X$, the feature extractor generates the feature matrix which is the input of the classifier ELM.

Next, we will introduce three important components of logistic-ELM, including feature extractor, logistic-ELM classifier, and pseudo-random sequence generator by logistic mapping.

#### 3.1 Component-1: logistic-ELM classifier

The proposed logistic-ELM classifier is the core component of the proposed method, as shown in Fig. 3. The purpose of this component is to classify faults based on the feature matrix $F$. Based on the original ELM, we generate the input weights through logistic mapping and omit the biases $b$ of the ELM with $L$ hidden neurons and activation function $g(\cdot)$, the logistic-ELM can be written as

$$\sum_{j=1}^L \beta_j g(w_j \cdot f_i) = c_i, \ i = 1, \ldots, N$$

(8)

where $w_j = [w_{j1}, w_{j2}, \ldots, w_{jk}]^T$ is the weight vector connecting the input feature $f_i$ and $j$th hidden neuron, $\beta_j = [\beta_{j1}, \beta_{j2}, \ldots, \beta_{jm}]^T$ is the weight vector connecting the $j$th hidden neuron and output $c_i$.

Let $W = [w_1, w_2, \ldots, w_L]$ be the input weights, and $\beta = [\beta_1, \beta_2, \ldots, \beta_L]^T$ be the output weights, such that, $H\beta = C$, where

$$H = \begin{bmatrix} g(w_1 \cdot f_1) & \cdots & g(w_L \cdot f_1) \\ \vdots & \ddots & \vdots \\ g(w_1 \cdot f_N) & \cdots & g(w_L \cdot f_N) \end{bmatrix}_{N \times L}$$

(9)

In the ideal situation, the factual output of the classifier is the same as the fault label, specifically, $C = Y$. Consequently, $H\beta = Y$, and $\hat{\beta} = H^+ Y$. The process is illustrated in Fig. 4.

#### 3.2 Component-2: pseudo-random sequence generator for input weights based on chaotic logistic mapping

As described previously, the ELM can randomly initialize the input weights and hidden layer biases to obtain the corresponding output weights, and thus, this method can be adapted to different applications. But random initialization may lead to a certain degree of uncertainty in the classification result. Therefore, we introduce the pseudo-random sequence generator of Logistic Mapping for input weights.
of ELM. In fact, this component is also a necessary part of logistic-ELM as shown in Fig. 3. Since the function of the input weights is to map the feature matrix to the hidden neurons, the correlation between the elements of the input weights matrix should be as small as possible. To this end, we generate the input weights through logistic mapping. Logistic mapping involves the stability and instability of the complete and part of the situation, respectively, and is
extremely appropriate to generate the initial input weights. The logistic mapping can be expressed as

$$z_k = \mu z_{k-1}(1 - z_{k-1})$$

(10)

where $z_1 \in (0, 1)$ and $\mu \in (3.56995, 4)$. Consequently, $z = (z_1, z_2, ...)$, and we can obtain the initial input weights $W$ by arranging the elements in the sequence $z$. Therefore, $W$ can be expressed as

$$W = \begin{bmatrix}
Z_1 & \cdots & Z_{K(L-1)+1} \\
\vdots & \ddots & \vdots \\
Z_K & \cdots & Z_{KL}
\end{bmatrix}$$

(11)

In this case, a high degree of irrelevance must be ensured for $W$, and thus, the biases do not need to be set. The biases can be assigned a zero value during the classification. The input weights are generated in Algorithm 1.

### Algorithm 1 Weight Generator $(z_1, \mu, L, K)$.

**Require:** $(z_1, \mu, L, K)$

**Ensure:** $W$

1: for $k$ 2 to $K \times L$ by 1 do
2: $z_k = \mu z_{k-1}(1 - z_{k-1})$
3: end for
4: for $i$ 1 to $K$ by 1 do
5: for $j$ 1 to $L$ by 1 do
6: $W_{ij} = z_i(i-1) \times L + j$
7: end for
8: end for
9: return $W$

3.3 Component-3: feature extractor based on the mechanical vibration principle

The purpose of this component is to extract the feature matrix from vibration signals, as shown in Fig. 5.

Firstly, we extract 14 time-domain features through the mechanical principle, to express a certain situation of the vibration signals from different perspective, including Mean Value, Standard Deviation, Variance, Peak-to-Peak Value, Square Root Amplitude, Average Amplitude, Mean Square Amplitude, Peak Value, Waveform Index, Peak Index, Impulsion Index, Clearance Factor, Degree of Skewness, and Kurtosis Value. Each feature represents one characteristic of the original vibration signals, and related equations are as follows.

![Fig. 5 The Process to Build the Feature Extractor](image-url)
• Mean value: \( \text{mean}(x_i) = \frac{1}{n} \sum_{t=1}^{n} x_{i,t} \)
• Standard deviation: \( \sigma(x_i) = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (x_{i,t} - \text{mean}(x_i))^2} \)
• Variance: \( \sigma^2(x_i) = \frac{1}{n} \sum_{t=1}^{n} (x_{i,t} - \text{mean}(x_i))^2 \)
• Peak-to-Peak Value: \( V_{pp}(x_i) = \max(x_i) - \min(x_i) \)
• Square root amplitude: \( X_r(x_i) = \frac{\sqrt{\sum_{t=1}^{n} (x_{i,t} - \text{mean}(x_i))^2}}{\text{mean}(x_i)} \)
• Average amplitude: \( \text{mean}(x_i) = \frac{1}{n} \sum_{t=1}^{n} x_{i,t} \)
• Mean square amplitude: \( X_{rms}(x_i) = \frac{\sum_{t=1}^{n} x_{i,t}^2}{\text{mean}(x_i)} \)
• Peak value: \( X_p(x_i) = \max(|\text{mean}(x_i)|) \)
• Waveform Index: \( X_w(x_i) = \frac{X_{rms}(x_i)}{|\text{mean}(x_i)|} \)
• Impulsion Index: \( C_f(x_i) = \frac{X_p(x_i)}{X_{rms}(x_i)} \)
• Clearance factor: \( C_e(x_i) = \frac{X_{rms}(x_i)}{\text{mean}(x_i)} \)
• Degree of skewness: \( C_w(x_i) = \frac{1}{n} \sum_{t=1}^{n} (x_{i,t} - \text{mean}(x_i))^3}{\sigma(x_i)^3} \)
• Kurtosis value: \( C_q(x_i) = \frac{1}{n} \sum_{t=1}^{n} (x_{i,t} - \text{mean}(x_i))^4}{\sigma(x_i)^4} \)

The above 14 feature models can be combined to create a feature pool \( S = (s_1, ..., s_{14}) \) of candidate features from which the best subset \( S' \) is selected by SFS procedure. Initializing the best subset as an empty set, the SFS starts the iterative processes to find appropriate features. The selection condition is of best accuracy by logistic-ELM component (described in Sect. 3.1). By comparing the classification result \( C \) with the fault label \( Y \), the accuracy can be calculated as

\[
\text{Accuracy}(C, Y) = \frac{\sum_{i=1}^{N} P(c_i = y_i) \times 100}{N}
\]

where \( P(c_i = y_i) = \begin{cases} 1, & c_i = y_i \\ 0, & c_i \neq y_i \end{cases} \)

Repeat the iteration until the accuracy of the best subset does not improve with the further addition of a feature. The feature extractor is built by the feature models in the best subset. Related SFS algorithm for feature selection is described in Algorithm 2.

**Algorithm 2 SFS \((S, Y, W, \beta)\).**

Require: \((S, Y, W, \beta)\)
Ensure: \(S'\)
1: \(S' = \emptyset, s_0 = \emptyset\)
2: \(\text{Acc} = -1, \text{sign} = 0, \text{max} = 0\)
3: while \(\text{Acc} < \text{max} \) do
4: \(\text{Acc} = \text{max}\)
5: add \(s_{\text{sign}}\) to \(S'\)
6: delete \(s_{\text{sign}}\) from \(S\)
7: for \(k = 1\) to card(\(S\)) by \(1\) do
8: \(C = g(S' \ast W) \ast \beta\)
9: if Accuracy\((C, Y) > \text{max}\) then
10: \(\text{max} = \text{Accuracy}\((C, Y)\)
11: \(\text{sign} = k\)
12: end if
13: end for
14: end while
15: return \(S'\)

Table 1 Sample description

| Type (inches) | Label | #Train | #verify | #Test |
|---------------|-------|--------|---------|-------|
| N 0           | 1     | 81926  | 20480   | 20480 |
| IF 0.007      | 2     | 81926  | 20480   | 20480 |
| IF 0.021      | 3     | 81926  | 20480   | 20480 |
| BF 0.007      | 4     | 81926  | 20480   | 20480 |
| BF 0.021      | 5     | 81926  | 20480   | 20480 |
| OF3 0.007     | 6     | 81926  | 20480   | 20480 |
| OF3 0.021     | 7     | 81926  | 20480   | 20480 |
| OF6 0.007     | 8     | 81926  | 20480   | 20480 |
| OF6 0.021     | 9     | 81926  | 20480   | 20480 |
| OF12 0.007    | 10    | 81926  | 20480   | 20480 |
| OF12 0.021    | 11    | 81926  | 20480   | 20480 |

Frequency=12 kHz; motor load=0 hp; speed=1797 r/min
4 Experiments and analysis

To demonstrate the effectiveness of the logistic-ELM, we use the rolling bearing vibration signal dataset [43] prepared by the CWRU bearing data centre to conduct experiments. This dataset is a world-recognized standard dataset for bearing fault diagnosis and contains a large number of rolling bearing vibration signals under normal and fault conditions. In this paper, we select bearing data with two kinds of fault diameter (0.007 inch and 0.021 inch) for classification. The motor load and speed are 0 hp and 1797 r/min, respectively. The sampling frequency is 12 kHz. There exist 6 fault types: normal (N), inner raceway fault (IF), ball fault (BF), and outer raceway faults located at 3 o’clock (OF3), 6 o’clock (OF6), and 12 o’clock (OF12). We select 60 original samples for each fault type, where each sample has 2048 sampling signal points. The sample description for the different fault types is presented in Table 1.

Figure 6 shows the original vibration signals of the 11 fault types, where the green and red lines indicate the normal signals and fault signals, respectively. The amplitude of the normal signal is small, as shown in Fig. 6a, and the periodicity is not notable. We observe that there are great amplitude deviations between fault signals and normal signals, as
shown in Fig. 6b–k, and some of the fault signals generally demonstrate an obvious periodicity. As shown in Fig. 6b, c, we can observe periodic in inner race fault with different sizes. The 0.007-inch inner race fault causes the amplitude to increase by approximately 4 times, while the 0.021-inch inner race fault causes the amplitude to increase by approximately 10 times. Intuitively, for inner race fault, larger diameter leads to larger amplitude. Theoretically, the inner race and the balls without any fault achieve a snug fit and run smoothly. Conversely, inner race fault can result in the gap between the inner race and the balls, as well as the vibration. Obviously, more serious fault widens the gap which causes larger amplitude of the vibration. The ball fault causes the amplitude to increase by approximately 2 times, and the signal does not exhibit any notable periodicity, as shown in Fig. 6d, e, likely because the balls roll constantly, and the contact points with the raceway change randomly. Coupled with the effect of the lubricating oil, the fault performance is not sufficiently clear. The outer raceway faults are periodic with a large amplitude as shown in Fig. 6f–k. For the outer raceway faults located at 3 o’clock and 12 o’clock, larger diameter leads to larger amplitude. On the contrary, for the outer raceway faults located at 6 o’clock, larger diameter leads to smaller amplitude. Therefore, the linear correlation between fault diameter and amplitude cannot be determined.

The vibration signals of different bearing faults have different features and exhibit a certain regularity, which is the theoretical basis for fault diagnoses. However, the features of the signals cannot be obtained through only manual observation, and thus, a scientific algorithm design must be implemented.

We deploy seven experiments to verify the accuracy and efficiency of proposed logistic-ELM with comparisons from multiple perspectives. Experiments 1-5 are the analysis of logistic-ELM itself, and Experiments 6 and 7 are the comparison with existing state-of-the-art methods. Experiments 1, 2, 3, 4 and 7 are performed on CWRU dataset with fault diameter \( d = 0.007" \) & 0.021" and load = 0hp as described above, while experiment 5 is performed on CWRU dataset with fault diameter \( d=0.007" \) & 0.021" and load = 0hp & 1hp & 2hp & 3hp. Experiment 6 is performed on CWRU database with diameters identical to the comparison methods.

The purpose of each experiment is as follows. **Experiment 1** is to select appropriate number of hidden neurons for logistic-ELM. **Experiment 2** is to select initial chaotic values \( z_1 \) and \( \mu \) for Logistic Mapping of logistic-ELM. **Experiment 3** shows the process of selecting the features for logistic-ELM by SFS. **Experiment 4** is about predictive accuracy comparison with original ELM. Parameters of Experiments 1-4 are set in Table 2. **Experiment 5** is to verify the performance of logistic-ELM under different operating conditions. **Experiment 6** is about predictive accuracy comparisons with existing fault diagnosis methods. **Experiment 7** is about runtime cost comparisons with existing fault diagnosis methods.

### 4.1 Experiment 1. Select appropriate activation function and the number of hidden neurons for logistic-ELM

The purpose of this experiment is to select the appropriate activation function and the number of hidden neurons for logistic-ELM. We deploy the experiment on CWRU dataset with fault diameter \( d = 0.007" \) and 0.021", and Fig. 7 shows the accuracies of logistic-ELM under Sigmoid, Sine,
Hardlim, Triangular, and Radial functions, respectively, with various number of hidden neurons.

Overall, the predictive accuracy of Sigmoid is higher than other four activation functions. For Sigmoid, when the number of hidden neurons ranges from 1 to 20, the predictive accuracy increases as neurons number get higher. When the number of hidden neurons ranges from 20 to 30, the predictive accuracy maintains steady, although has little different. When the number of hidden neurons is bigger than 30, the predictive accuracy is 1, but algorithms may over-fit the training set now.

In order to scientifically describe the above phenomena, Pearson correlation coefficient \([44]\) is used to calculate the linear correlation between prediction accuracy and the number of hidden neurons, as

\[
\rho(\text{accuracy, neuron}) = \frac{\sigma_{\text{accuracy, neuron}}}{\sigma_{\text{accuracy}} \sigma_{\text{neuron}}}
\]  

(13)

where \(\sigma_{\text{accuracy, neuron}}\) is the covariance between the predictive accuracy and the number of hidden neurons. \(\sigma_{\text{accuracy}}\) and \(\sigma_{\text{neuron}}\) are variances of the predictive accuracy and the number of hidden neurons, respectively. Pearson correlation coefficient returns a value between \(-1\) and 1. A higher absolute value of the correlation coefficient indicates a stronger relationship between variables.

When the number of hidden neurons ranges from 1 to 20, Pearson correlation coefficient between prediction accuracy and the number of hidden neurons is 0.833, which indicates a strong positive correlation. When the number of hidden neurons ranges from 20 to 30, Pearson correlation coefficient is 0.116, which indicates a weak positive correlation. Therefore, we choose Sigmoid as the activation function and 20 as the number of hidden neurons.

### 4.2 Experiment 2. Select initial chaotic values for logistic mapping

This experiment is to select initial chaotic values for logistic-ELM. We deploy the experiment under different initial chaotic value of \(z_1\) and \(\mu\) generated through logistic mapping. Table 3 shows the predictive accuracies of logistic-ELM on CWRU dataset with fault diameter \(d = 0.007''\) and \(0.021''\). Logistic mapping is sensitive to initial conditions, and

Table 3 Predictive accuracy of logistic-ELM under different initial value of \(\mu\) and \(z_1\)

| \(z_1\) | 3.95 | 3.96 | 3.97 | 3.98 | 3.99 |
|---|---|---|---|---|---|
| 0.1 | 1 | 1 | 0.98 | 0.99 | 1 |
| 0.2 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| 0.3 | 1 | 1 | 0.99 | 0.99 | 0.99 |
| 0.4 | 1 | 0.98 | 0.99 | 0.98 | 0.99 |
| 0.5 | 1 | 0.98 | 0.98 | 0.98 | 0.99 |
| 0.6 | 1 | 0.99 | 0.96 | 1 | 1 |
| 0.7 | 0.98 | 1 | 0.99 | 1 | 1 |
| 0.8 | 0.98 | 0.98 | 1 | 0.96 | 1 |
| 0.9 | 0.99 | 0.98 | 0.99 | 0.99 | 0.99 |

Table 4 Process of feature selection under SFS

| Initial feature pool | Round 1 | Round 2 | Round 3 | Round 4 |
|---|---|---|---|---|
| Initial feature pool | \(\emptyset\) | \{F7\} | \{F7,F4\} | \{F7,F4,F9\} |
| Feature 1 | 0.19 | 0.72 | 0.95 | 0.96 |
| Feature 2 | 0.76 | 0.80 | 0.93 | 0.99 |
| Feature 3 | 0.48 | 0.86 | 0.93 | 0.98 |
| Feature 4 | 0.70 | 0.92 | - | - |
| Feature 5 | 0.59 | 0.87 | 0.93 | 0.99 |
| Feature 6 | 0.72 | 0.89 | 0.95 | 1.00 |
| Feature 7 | 0.79 | - | - | - |
| Feature 8 | 0.63 | 0.87 | 0.91 | 0.96 |
| Feature 9 | 0.58 | 0.92 | 0.98 | - |
| Feature 10 | 0.54 | 0.83 | 0.95 | 0.98 |
| Feature 11 | 0.58 | 0.80 | 0.88 | 0.96 |
| Feature 12 | 0.58 | 0.77 | 0.84 | 0.95 |
| Feature 13 | 0.43 | 0.70 | 0.91 | 0.96 |
| Feature 14 | 0.75 | 0.85 | 0.97 | 0.97 |

| Selected Feature | Feature 7 | Feature 4 | Feature 9 | Feature 6 |
|---|---|---|---|---|
| Feature 7 | 0.92 | \{F7\} | \{F7,F4\} | \{F7,F4,F9,F6\} |

| Table 5 Predictive accuracy comparison between original ELM and logistic-ELM |
|---|---|---|
| Prediction | Probability density |
| Accuracy | Original ELM | Logistic-ELM |
| 1 | 0.18 | 0.56 |
| [0.99,1) | 0.36 | 0.19 |
| [0.98,0.99) | 0.06 | 0.19 |
| [0.97,0.98) | 0.1 | 0 |
| [0.96,0.97) | 0.08 | 0.06 |
| [0.95,0.96) | 0.06 | 0 |
| [0.94,0.95) | 0.02 | 0 |
| [0.93,0.94) | 0.04 | 0 |
| [0.92,0.93) | 0 | 0 |
| [0.91,0.92) | 0.02 | 0 |
| [0.90,0.91) | 0 | 0 |
| [0.90,0.90) | 0.06 | 0 |
| Highest | 1 | 1 |
| Expected | 0.971 | 0.993 |
the closer $\mu$ is to 4, the more chaotic the system is. Therefore, let the ranges of initial values be $z_1 \in [0.1, 0.9]$ and $\mu \in [3.95, 3.99]$.

As shown in Fig. 4, when one of $z_1$ and $\mu$ is constant, there is no correlation between the other one and the prediction accuracy. Therefore, without considering the correlation, we choose $z_1$ and $\mu$ with a prediction accuracy of 1 in the table as the initial values of the logical mapping.

### 4.3 Experiment 3. Select appropriate time-domain features for logistic-ELM

This experiment intends to show the process of feature selection under Sequence Forward Selection (SFS). We execute our scheme with different combination of features by several rounds.

As shown in Table 4, the initial feature pool is empty in the first round and we run our scheme with each of the 14 features, respectively. Then we record the results and find that Feature 7 makes the highest accuracy 0.92 which is higher than 0.79. So we put Feature 4 in the feature pool. Now the feature pool has two elements. Then we repeat the operations above until the accuracy do not increase anymore. In this way, we obtain the feature pool of [Feature 7, Feature 4, Feature 9, Feature 6].

### 4.4 Experiment 4. Predictive accuracy comparison with original ELM

In this paper, logistic mapping is used to generate the input weights only once, and these weights can be directly used in the subsequent classification. Due to the chaotic nature of logistic mapping, we do not need to set biases while the original ELM randomly generates the input weights and biases to fit the generality. This experiment is to verify the efficiency and stability of logistic-ELM comparing with original ELM.

In order to observe the accuracy and stability of logistic-ELM versus the original ELM, we conducted several tests on each of them on the CWRU dataset with fault diameter $d = 0.007"$ and 0.021". The input weight matrix $W$ of the logistic-ELM is generated by the logistic mapping using the initial values chosen in Experiment 2 and bias $b$ is omitted. In original ELM, $W$ and $b$ are generated randomly. Other parameters are identical for both methods. As shown in table 5, the expected value of the original ELM prediction accuracy is 0.971, which is lower than the expected value of logistic-ELM of 0.993, although the highest prediction accuracy of both methods is 1. Moreover, the accuracy distribution of original ELM is scattered while logistic-ELM is stable. As shown in Fig. 8, the red line indicates logistic-ELM and the blue line indicates the original ELM. The horizontal coordinate of each point in the figure is the prediction accuracy and the vertical coordinate indicates the probability density for each prediction accuracy. Observe that the probability density of logistic-ELM is higher than it of the original ELM at each accuracy. Therefore, logistic-ELM is more stable and has higher prediction accuracy than the original ELM.

### 4.5 Experiment 5. Predictive accuracy under different operating conditions

The above-mentioned experiments demonstrate that the logistic-ELM is a fast diagnosis method with high predictive accuracy, but are all deployed under 0 hp motor load. In order to verify the logistic-ELM’s adaptability to different operating conditions, we perform an experiment to observe the predictive accuracy of fault diagnosis of the rolling bearings under four different loads (0 hp, 1hp, 2 hp, and 3 hp) selected from CWRU dataset. Table 6 shows the results under different operating conditions. We find that the highest predictive accuracy of logistic-ELM is 100% in these conditions.
conditions. It indicates that the logistic-ELM can adapt to variable operating conditions.

### 4.6 Experiment 6. Predictive accuracy comparisons with existing fault diagnosis methods

In this experiment, we select four existing fault diagnosis methods of rolling bearings for comparisons with proposed logistic-ELM, including DSLS-SVM [45], PMSEn [12], FuzzyMEn [13], VAEGAN-DRA [46], RCFOA-ELM [32] and DE-ELM [33], which were proposed from 2019 to 2021. Among these methods, DSLS-SVM is an extended algorithm of the SVM. PMSEn and FuzzyMEn are extended algorithms of the multi-scale entropy, while VAEGAN-DRA utilizes the deep learning method, generative adversarial network (GAN). RCFOA-ELM and DE-ELM are based on ELM. Datasets in these methods are subsets of CWRU but different, so we perform each comparison experiment on the same subset as described in the comparison object, with fivefold cross-validation. Table 7 shows the details of related sub datasets.

Due to the lack of understanding of the implementation details of these methods, the accuracies of related comparison objects are from their public reports. Figure 9 shows the comparisons with predictive accuracy of logistic-ELM. On the four sub datasets, our proposed logistic-ELM obtains accuracies with 100%, 99.71%, 98%, 100%, 100%, and 100%, respectively, while those of the four comparison objects are 99.9%, 94.89%, 97.27%, 100%, 98.34%,

### Table 7 Predictive accuracy comparison between original ELM and logistic-ELM

| Sub dataset                      | Description of fault information                                                                 |
|----------------------------------|--------------------------------------------------------------------------------------------------|
| Dataset-1 in DSLS-SVM [45]       | N(0 hp), IF-BF-OF(0 hp, 0.007") from CWRU dataset                                                |
| Dataset-2 in PMSEn [12]          | N(0 hp), N(1 hp), N(2 hp), N(3 hp), IF-BF-OF(0 hp, 0.021") from CWRU dataset                   |
| Dataset-3 in VAEGAN-DRA [46]     | N(0 hp), IF-BF-OF(0 hp, 0.007"), IF-BF-OF(0 hp, 0.014"), IF-BF-OF(0 hp, 0.021") from CWRU dataset |
| Dataset-4 in FuzzyMEn [13]       | N(0 hp), IF-BF-OF(0 hp, 0.007"), IF-BF(0 hp, 0.021"), OF(0 hp, 0.028") from CWRU dataset        |
| Dataset-5 in RCFOA-ELM [32]      | N(2 hp), IF-BF-OF(0 hp, 0.014") from CWRU dataset                                                |
| Dataset-6 in DE-ELM [33]         | N(2 hp), IF-BF-OF(2 hp, 0.007") from CWRU dataset                                                |

### Table 8 Runtime cost comparisons (s)

| Method         | Average runtime Cost (s) |
|----------------|--------------------------|
| DSLS-LSSVM     | 0.8311                   |
| PMSEn          | 73.04                    |
| FS-CNN         | 3.586                    |
| Proposed       | 0.0393                   |
and 99.5%. Results indicate the proposed logistic-ELM is accurate in fault diagnosis and has superiority to existing methods.

4.7 Experiment 7. Runtime cost comparisons with existing fault diagnosis methods

In order to assess the runtime cost of logistic-ELM, the comparison are made with existing fault diagnosis methods. We select DSLS-LSSVM [45], PMSEn [12] and FS-CNN [47] for comparisons, which are based on SVM, multi-scale entropy, and CNN, on CWRU dataset with fault diameter \( d = 0.007\)”. All the codes are implemented in MATLAB R2018b, in operating system Win10 (64-bit), with a 16G Memory and Intel (R) Core (TM) i7-10875H CPU. To obtain the calculation time in a reasonable manner, the experiments are performed 5 times, and the average time is determined. The dataset for each experiment includes 300 samples, and the length of each sample is 2048. Table 8 shows the results, where runtime is just testing time, not training time. As we can see, the average predicting runtime cost is 0.0393 s, while the other three methods need 0.8311 s, 73.04 s, and 3.586 s, respectively. Compared with the existing method, the operating efficiency of logistic-ELM is increased by 21.15 times, 1858.52 times, and 9.12 times, respectively. Results indicate the proposed logistic-ELM is a very fast fault diagnosis method for rolling bearings.

5 Conclusion

This paper proposes a fast and efficient fault diagnosis method for rolling bearings, named logistic-ELM. We obtain optimal features by SFS and improve the stability of our method by combining original ELM with logistic mapping. Indicated by the experiments, the key conclusions of the paper are as follows:

- The proposed logistic-ELM has great high predictive accuracy and very short runtime.
- Logistic-ELM can adapt to different operating conditions.
- During all comparisons, the logistic-ELM presents high superiority in fault diagnosis of rolling bearings.

In future work, we will research on compound faults and prognostic and health management (PHM).

Acknowledgements This work was funded by the National Key Research and Development Program of China under Grant No. 2019YFB1405803.

Author contributions Zhenhua Tan proposed the original opinion, methodology, design, original architecture and writing. Jingyu Ning is responsible for the design, coding, data analysis, original writing, and revision. Kai peng participated in the coding, data analysis, and revision, while Zhenche Xia and Danke Wu participated in the design and revision.

References

1. Sheng S (2016) Wind turbine gearbox reliability database, condition monitoring. https://www.nrel.gov/docs/ fy16osti/66028.pdf
2. Wheeler PG (1968) Bearing analysis keeps downtime down. Plant Eng 25:87–89
3. Martin RL (1970) Detection of ball bearing malfunctions. Instrum Control Syst 43(12):79–82
4. Boto PA (1971) Detection of bearing damage by shock pulse measurement. Ball Bearing J 54(4):12
5. Harting DC, Taylor JW (1974) Demodulated resonance analysis system. US Patent 3,842,663. https://patentimages.storage.googleapis.com/pdfs/US3842663.pdf
6. Cooley JW, Tukey JW (1965) An algorithm for the machine calculation of complex Fourier series. Math Comput 19:297–301. https://doi.org/10.1090/S0025-5718-1965-0178586-1
7. Nikula R-P, Karioja K, Pylvänäinen M, Leiviskä K (2020) Auto-regression of low-speed bearing fault diagnosis based on autocorrelation of time domain features. Mech Syst Signal Process 138:106572. https://doi.org/10.1016/j.ymssp.2019.106572
8. Richman JS, Moorman JR (2000) Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 278(6):2039–2049. https://doi.org/10.1152/ajpheart.2000.278.6.H2039 (PMID: 10843903)
9. Pincus SM (1991) Approximate entropy as a measure of system complexity. Proc Natl Acad Sci 88(6):2297–2301. https://doi.org/10.1073/pnas.88.6.2297
10. Li X, Cui S, Voss L (2008) Using permutation entropy to measure the electroencephalographic effects of Sevoflurane. Anesthesiology 109(3):448–456. http://doi.org/10.1097/ALN.0b013e318182a91b
11. Nicolaou N, Georgiou J (2012) Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. Expert Syst Appl 39(1):202–209. https://doi.org/10.1016/j.eswa.2011.07.008
12. Zhao D, Liu S, Cheng S, Sun X, Wang L, Wei Y, Zhang H (2021) Parallel multi-scale entropy and its application in rolling bearing fault diagnosis. Measurements 168:108333. https://doi.org/10.1016/j.measurement.2020.108333
13. Howedi A, Lotfi A, Pourabdollah A (2020) A multi-scale fuzzy entropy measure for anomaly detection in activities of daily living. In: Proceedings of the 13th ACM international conference on pervasive technologies related to assistive environments. PETRA ’20. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3389189.3397987
14. Stack JR, Harley RG, Habetler TG (2004) An amplitude modulation detector for fault diagnosis in rolling element bearings. IEEE Trans Indus Electron 51:1097–1102
15. Blankenship GW, Singh R (1995) Analytical solution for modulation sidebands associated with a class of mechanical oscillators, J Sound Vib 179(1):13–36. https://doi.org/10.1006/jsvi.1995.0002
16. Liang B, Iwinski SD, Zhao Y (2013) Application of power spectrum, cepstrum, higher order spectrum and neural network analyses for induction motor fault diagnosis. Mech Syst Signal Process 39(1):342–360. https://doi.org/10.1016/j.ymssp.2013.02.016
17. Feng Z, Liang M, Chu F (2013) Recent advances in time-frequency analysis methods for machinery fault diagnosis: a review with application examples. Mech Syst Signal Process 38(1):165–205. https://doi.org/10.1016/j.ymssp.2013.01.017
18. Parker B Jr, Ware HA, Wipf DP, Tompkins WR, Clark BR, Larson EC, Poor HV (2000) Fault diagnostics using statistical change detection in the bispectral domain. Mech Syst Signal Process 14(4):561–570. https://doi.org/10.1006/mssp.2000.1299
19. Tian J, Morillo C, Azarian M, Pecht M (2015) Motor bearing fault detection using spectral kurtosis based feature extraction and k-nearest neighbor distance analysis. IEEE Trans Industr Electron 63:1–1. https://doi.org/10.1109/TIE.2015.2509913
20. Antoni J, Randall RB (2006) The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines. Mech Syst Signal Process 20(2):308–331. https://doi.org/10.1016/j.ymssp.2004.09.002
21. Vapnik VN (2000) The nature of statistical learning theory. In: Statistics for engineering and information science. Springer, New York. https://doi.org/10.1007/978-1-4757-3264-1
22. Konar P (2011) Chattopadhyay P Bearing fault detection of induction motor using wavelet and support vector machines (svms). Appl Soft Comput 11(6):4203–4211. https://doi.org/10.1016/j.asoc.2011.03.014
23. Mian T, Choudhary A (2022) Fatima S A sensor fusion based approach for bearing fault diagnosis of rotating machine. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 236(5):661–675. https://doi.org/10.1177/174806X211044831
24. Choudhary A, Goyal D, Letha SS (2021) Infrared thermography-based fault diagnosis of induction motor bearings using machine learning. IEEE Sens J 21(2):1727–1734. https://doi.org/10.1109/JSEN.2020.3015868
25. Mishra R.K, Choudhary A, Mohanty A.R, Fatima S (2021) Multi-domain bearing fault diagnosis using support vector machine. In: 2021 IEEE 4th international conference on computing, power and communication technologies (GUCon), pp 1–6. https://doi.org/10.1109/GUCon50781.2021.9573613
26. Mian T, Choudhary A, Fatima S (2022) Multi-sensor fault diagnosis for misalignment and unbalance detection using machine learning. In: 2022 IEEE international conference on power electronics, smart grid, and renewable energy (PESGRE), pp. 1–6. https://doi.org/10.1109/PESGRE52268.2022.9715938
27. Muralidharan V, Sugumaran V (2012) A comparative study of naive Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. Appl Soft Compt 12(8):2023–2029. https://doi.org/10.1016/j.asoc.2012.03.021
28. Lecun Y, Bengio Y (1995) Convolutional networks for images, speech, and time-series. In: The handbook of brain theory and neural networks. MIT Press, Cambridge, MA, USA, pp 255–258. ISBN: 0262511026. https://doi.org/10.5555/303568.303704
29. Huang NE, Zhu Q-Y, Shih HH, Zheng Q, Yen N-C, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond Ser A 454(1971):903–998. https://doi.org/10.1098/rspa.1998.0193
30. Smith JR (2005) The local mean decomposition and its application to EEG perception data. J R Soc Interface 2:443–454
31. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536. https://doi.org/10.1038/323533a0
32. He C, Wu T, Gu R, Jin Z, Ma R, Qu H (2021) Rolling bearing fault diagnosis based on composite multiscale permutation entropy and reverse cognitive fruit fly optimization algorithm—extreme learning machine. Measurement 173:108636. https://doi.org/10.1016/j.measurement.2020.108636
33. Hu Y, Dong M, Wang G, Fan Z, Zhang S (2019) New method of bearing fault diagnosis based on mmemd and de_elm. J Eng 2019(23):9152–9156. https://doi.org/10.1049/joe.2018.9206
34. Choudhary A, Mian T, Fatima S, Panigrahi BK (2022) Passive thermography based bearing fault diagnosis using transfer learning with varying working conditions. IEEE Sens J. https://doi.org/10.1109/JSEN.2022.3164430
35. Choudhary A, Mian T, Fatima S (2021) Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images. Measurement 176:109196. https://doi.org/10.1016/j.measurement.2021.109196
36. May R (1976) Simple mathematical models with very complicated dynamics. Nature 266:457–459. https://doi.org/10.1038/261459a0
37. Li Y, Gao Q, Miao B, Zhang W, Liu J, Zhu Y (2021) Application of the refined multiscale permutation entropy method to fault detection of rolling bearing. J Braz Soc Mech Sci Eng. https://doi.org/10.1007/s40430-021-02986-7
38. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen N-C, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond Ser A 454(1971):903–998. https://doi.org/10.1098/rspa.1998.0193
39. Smith JR (2005) The local mean decomposition and its application to EEG perception data. J R Soc Interface 2:443–454
40. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536. https://doi.org/10.1038/323533a0
41. Han T, Tian Z, Yin Z, Tan A (2020) Bearing fault identification based on convolutional neural network by different input modes. J Braz Soc Mech Sci Eng. https://doi.org/10.1007/s40430-020-02561-6
42. Luo M, Li C, Zhang X, Li R, An X (2016) Compound feature selection and parameter optimization of elm for fault diagnosis of rolling element bearings. ISA Trans 65:556–566. https://doi.org/10.1016/j.isatra.2016.08.022
43. KA L (2012) Case western reserve university bearing data center. https://engineering.case.edu/bearingdatacenter/download-data-file
44. Profillidis VA, Botzoris GN (2019) Statistical methods for transport demand modeling. Model Transp Demand. https://doi.org/10.1109/B978-0-12-811513-8.00005-4
45. Li X, Yang J, Pan H, Cheng J, Cheng J (2019) A novel deep stacking least squares support vector machine for rolling bearing fault diagnosis. Comput Ind 110:36–47. https://doi.org/10.1016/j.compind.2019.05.005
46. Liu S, Jiang H, Wu Z, Li X (2021) Rolling bearing fault diagnosis using variational autoencoding generative adversarial networks with deep regret analysis. Measurement 168:108371. https://doi.org/10.1016/j.measurement.2020.108371
47. Liang P, Deng C, Wu J, Yang Z, Zhu J (2019) Intelligent fault diagnosis of rolling element bearing based on convolutional neural network and frequency spectrograms. In: 2019 IEEE international conference on prognostics and health management (ICPHM), pp 1–5. https://doi.org/10.1109/ICPHM.2019.8819444