An effective multi-channel fault diagnosis approach for rotating machinery based on multivariate generalized refined composite multi-scale sample entropy

Zhenya Wang · Hui Chen · Ligang Yao · Xu Chen · Xiaoli Qi · Jun Zhang

Received: 12 January 2021 / Accepted: 12 August 2021 / Published online: 16 September 2021 © The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract Fault diagnosis of critical rotating machinery components is necessary to ensure safe operation. However, the commonly used rotating machinery fault diagnosis methods are generally based on the single-channel signal processing method, which is not suitable for processing multi-channel signals. Thus, to extract features and carry out the intelligent diagnosis of multi-channel signals, a novel method for rotating machinery fault diagnosis is proposed. Firstly, a novel nonlinear dynamics technique named the multivariate generalized refined composite multi-scale sample entropy was presented and applied to extract fusion entropy features of multi-channel signals. Secondly, a practical manifold learning known as supervised isometric mapping was introduced to map the high-dimensional fusion entropy features in a low-dimensional space. In a third step, the Harris hawks optimization-based support vector machine was applied to carry out the intelligent fault recognition. Finally, aiming to verify the effectiveness of the proposed method and present its advantages, it was applied to analyze the rotating machinery system bearing and gear data. The experimental results have shown that the method at hand can accurately identify various faults in both the bearings and gears. Furthermore, in addition to being suitable for multi-channel signal fault diagnosis, it had higher recognition accuracy compared to other multi-channel or single-channel methods.

Keywords Fault diagnosis · Rotating machine · Multi-scale entropy · Isometric mapping · Support vector machine

1 Introduction

Rotating machinery has been widely employed in various industrial fields, including petroleum, aerospace, metallurgy, energy, and chemistry. As critical rotating machinery components, bearings and gears have a significant role when it comes to supporting shafts, power conversion, and power transmission. However, due to complex operating conditions, their performance gradually degrades during a long-term operation and faults occur. Such behavior affects the performance of rotating machinery, their reliability,
and service life, often causing inevitable economic losses. Thus, it is necessary to implement an effective intelligent fault diagnosis method for both bearings and gears [1–3].

1.1 Motivation

During the previous two decades, intelligent fault diagnosis methods based on vibration signals were gradually developed and applied in rotating machinery. Such methods were commonly used as they were proved to be effective [4]. They mostly included two steps: (1) the signal processing algorithms were firstly used to extract fault features and (2) various pattern recognition algorithms were applied to diagnose the faults based on the extracted features [5]. Commonly used intelligent fault diagnosis methods were often based on single-channel signals [6–10].

However, for rotating machinery, a single information source often contains limited fault information, meaning that it may result in ambiguous or even plain wrong diagnosis. Additionally, with the continuous improvement of rotating machinery safety requirements, this type of technology cannot meet the growing demand for data and information mining [11]. Thus, the fusion data processing methods based on multi-channel signals attracted the attention of scholars aiming to detect and diagnose the situation in detail. Such multi-channel signal processing methods are slowly becoming a standard part of intelligent fault diagnosis development [12, 13]. For these technologies, signals obtained by the multiple sensors are processed as a whole, enabling them to fully extract the associated information, improve the information utilization rate, and finally obtain more accurate and reliable diagnosis results.

Moreover, the fusion features obtained using the multi-channel rotating machinery signals are regularly high-dimensional, nonlinear, and mixed with redundancy. Such properties increase the recognition time and reduce the diagnostic accuracy of subsequent classifiers [14]. When it comes to selecting low-dimensional and sensitive feature sets from these massive and nonlinear fusion features, it is imperative to consider intelligent fault diagnosis methods.

1.2 Entropy-based feature extraction approach

The rotating machinery vibration signals are usually nonlinear and non-stationary, which significantly increases the fault feature extraction complexity [14]. Fortunately, with the continuous development of nonlinear theory, entropy-based feature extraction algorithms which can overcome this problem were developed, as mentioned earlier. As such, they are widely applied for the feature extraction of various mechanical signals. The said technology mainly includes approximate entropy, sample entropy, permutation entropy, fuzzy entropy, and dynamic symbolic entropy [15–19].

Sample entropy is a frequently used method, which can effectively detect the minuscule time series changes. When the rotating machinery faults occur, the corresponding nonlinear dynamics complexity changes accordingly, requiring the application of sample entropy-based feature extraction algorithms [20–23]. For example, Wang combined the sample entropy method with ensemble empirical mode decomposition to identify the rolling bearing faults [20]. Gao et al. combined the multi-scale sample entropy with local mean decomposition to effectively extract the wind turbine fault features [21]. Furthermore, Dai et al. proposed a composite multi-scale sample entropy algorithm developed by combining the composite coarse-grained structure and the multi-scale sample entropy. The algorithm was proved to effectively extract the bearing features [22]. Recently, Wang et al. have proposed the generalized refined composite multi-scale sample entropy algorithm and verified its performance by using bearing data [23].

The above-mentioned sample entropy-based algorithms performed very well in rotating machinery fault diagnosis; however, all the algorithms belong to the group of single-channel analysis methods. Since such methods are not suitable for processing multi-channel signals, it is vital to explore the state-of-the-art multi-channel signal processing methods. Lu and Wang proposed the multivariate multi-scale sample entropy (MMSE) algorithm based on multi-scale sample entropy [24]. The algorithm was used to analyze any number of channel signals, providing reliable complexity analysis results. Nevertheless, when using MMSE to extract the fusion features of multi-channel rotating machinery signals, two shortcomings remained. (1) The new multivariate coarse-grained
structure (MCS) of MMSE only considers the single MCS under the scale factor $s$, ignoring other equally important MCSs. (2) MMSE obtains a new MCS by sliding a window and averaging all the data within, which can cause the amplitude information of multi-channel signals to be ignored [25, 26].

However, a number of studies were carried out aiming to improve the MMSE. To solve the first shortcoming, Humeau-Heurtier proposed the multivariate refined composite multi-scale sample entropy (MRCMSE) algorithm. The algorithm considered multiple MCSs (i.e., multivariate composite coarse-grained structures) under $s$. Furthermore, the refined operation was used to avoid invalid entropy values [25]. Additionally, regarding the second deficiency, Yin et al. introduced the multivariate generalized multi-scale sample entropy (MGCMSE). The original average calculation method applied in MMSE was replaced by variance operation in the MCS construction of MMSE. The authors demonstrated the MGCMSE effectiveness by distinguishing traffic detectors of various complexities [26].

Unfortunately, both the MRCMSE and the MGCMSE only solve one MMSE shortcoming. Thus, a novel MGRCMSE algorithm was proposed in this paper, aiming to overcome both defects by combining MRCMSE and MGCMSE. To prove its effectiveness, the algorithm was applied to extract the fusion feature of multi-channel vibration signals.

1.3 The proposed multi-channel fault diagnosis method

In this study, the authors have proposed an intelligent fault diagnosis method based on the fusion entropy manifold features and optimized support vector machine. The proposed method enables the effective mining of the fusion features from multi-channel rotating machinery signals, increasing the fault diagnosis accuracy, which is the main contribution of this paper. The method is divided into three phases:

- **The first phase**—a novel nonlinear algorithm, named multivariate generalized refined composite multi-scale sample entropy (MGRCMSE), was developed to enable mining the fusion features from multi-channel signals.
- **The second phase**—includes utilizing the supervised isometric mapping (SIM) [23] algorithm to reduce the fusion entropy feature dimensionality, enabling the authors to obtain the low-dimensional and sensitive fusion entropy manifold feature sets.
- **The third phase**—is focused on fault identification and diagnosis; the Harris hawks optimization-based support vector machine (HHO-SVM) [27] was applied to both identify and diagnose the faults.

The effectiveness of the proposed method was verified by analyzing the fault diagnosis experiments of bearings and gears mounted within the rotating machinery system. Moreover, compared to other single-channel or multi-channel feature extraction methods, the proposed feature extraction method extracted features from the original signals more effectively and with a higher fault diagnosis ability.

1.4 Contributions and structure of this paper

The main contributions of this paper can be stated as follows:

- A novel nonlinear dynamics analysis algorithm (MGRCMSE) was proposed.
- The proposed algorithm was used to carry out the multi-channel fault diagnosis of the rotating machinery.
- The proposed algorithm was combined with the SIM to enable sensitive feature extraction.

Finally, the paper was structured as follows: Sect. 2 outlines the mathematical background behind the MGRCMSE. The novel intelligent multi-channel fault diagnosis method for rotating machinery is presented in Sect. 3. The bearings and gears were used as test specimens in the experimental verification (Sect. 4). Finally, conclusions are given in Sect. 5.

2 Proposed multivariate generalized refined composite multi-scale sample entropy approach

The MMSE and MGCMSE algorithms were introduced in this section. Additionally, based on these two algorithms, the MGRCMSE was proposed after implementing the multivariate generalized composite coarse-grained structure and refined operation. The flowchart of these three algorithms allowing the easier comparison is shown in Fig. 1.
2.1 Multivariate multi-scale sample entropy

For a multi-channel time series $X_K = \{x_{k,j}\}_{j=1}^{N}$, $K = 1, 2, ..., p$ is the number of channels, and $N$ is the number of sample points of each channel time series. The specific MMSE steps are as follows:

1. The MCSs $Y_k^s = \{y_{k,j}^s\}_{j=1}^{Ns}$ of multi-channel signals at $s$ are constructed using Eq. (1), where $Ns = \lfloor N/s \rfloor$:

$$y_{k,j}^s = \frac{1}{s} \sum_{i=(j-1)s+1}^{is} x_{k,i}$$

(1)

2. According to the time delay embedding theory, the corresponding delay vector $Y_m^s(j)$ is found for each $Y_k^s$:

$$Y_m^s(j) = \{y_{1,j}^s, y_{1,j+1}, \ldots, y_{1,j+(m_1-1)\tau_1},$$
$$y_{2,j}^s, y_{2,j+1}, \ldots, y_{2,j+(m_2-1)\tau_2}, \ldots, y_{p,j}^s,$$
$$y_{p,j+\tau_p}, \ldots, y_{p,j+(m_p-1)\tau_p}\}$$

(2)

where $M = [m_1, m_2, \ldots, m_p]$ is the embedding vector, $\tau = [\tau_1, \tau_2, \ldots, \tau_p]$ is the delay vector, $j = 1, 2, \ldots, Ns - L$, and $L = \max(M) \times \max(\tau)$.

3. The maximum norm value of any two delay vectors $Y_m^s(j)$ and $Y_m^s(h)$ is found as the distance between them:

$$d[Y_m^s(j), Y_m^s(h)] = \max_{l=1, \ldots, m} \{|y_l^s(j) - y_l^s(h) + l - 1)|\}$$

(3)

If $d[Y_m^s(j), Y_m^s(h)] \leq r, j \neq h$, these two vectors are similar, with $r$ being a similarity tolerance.

4. Variables $B_{m,j}^s = d[Y_m^s(j), Y_m^s(h)] \leq r$ and $P_{m,j}(r) = B_{m,j}^s/(Ns - L - 1)$ are defined. Then, the similarity probability of any two delay vectors at the embedding dimension $m$ can be calculated as:

$$P_m^s(r) = \frac{1}{Ns - L} \sum_{j=1}^{Ns-L} P_{m,j}(r)$$

(4)

5. Steps (1) to (4) are repeated to calculate the similarity probability of any two delay vectors at $m + 1$ dimension. To ensure that the system is completely transformed into $m + 1$ dimension, the embedding size of other channels remains the same. Next, $B_{m+1,j}^s = d[Y_{m+1}^s(j), Y_{m+1}^s(h)] \leq r$ and $P_{m+1,j}(r) = B_{m+1,j}^s/[p(Ns - L - 1)]$ are
defined, followed by the calculation of the similarity probability of any two delay vectors at the embedding dimension \( m + 1 \) using Eq. (5):

\[
P_{m+1}^s(r) = \frac{1}{p(Ns-L)} \sum_{j=1}^{p(Ns-L)} P_{m+1,j}^s(r) \quad (5)
\]

6. The final MMSE values can be computed by:

\[
\text{MMSE}(X_K, s, m, r, N) = - \ln \left[ \frac{P_{m+1}^s(r)}{P_{m}^s(r)} \right] \quad (6)
\]

2.2 Multivariate generalized multi-scale sample entropy

In the MGMSE algorithm, the variance calculation replaces the average calculation used in MMSE, retaining the time series amplitude information. For the multi-channel time series \( X_K = \{x_{K,j}\}_{j=1}^N \), the main MGMSE steps can be summarized as follows:

1. The multivariate generalized coarse-grained structures \( Y_{G,K}^s = \{y_{G,K,j}^s\}_{j=1}^{Ns} \) of multi-channel signals at \( s \) are constructed:

\[
y_{G,K,j}^s = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} (x_{K,i} - \bar{x}_{K,i})^2 \quad (7)
\]

\[
\bar{x}_{K,i} = \frac{1}{s} \sum_{i=(j-1)s+1}^{js} x_{K,i}
\]

2. Similarly, for each \( Y_{G,K}^s \), the similarity probabilities \( P_{G,m}^s(r) \) and \( P_{G,m+1}^s(r) \) can be calculated by applying the MMSE steps (2) to (5).

3. Lastly, the MGMSE output values can be computed as:

\[
\text{MGMSE}(X_K, s, m, r, N) = - \ln \left[ \frac{P_{m+1}^s(r)}{P_{m}^s(r)} \right] \quad (8)
\]

2.3 Multivariate generalized refined composite multi-scale sample entropy

Based on MGMSE, the MGRCMSE algorithm was proposed in this paper, which was developed combining the multivariate composite coarse-grained structure and refined entropy operation. For \( X_K = \{x_{K,j}\}_{j=1}^N \), the key MGRCMSE steps are:

1. The multivariate generalized composite coarse-grained structures \( Y_{G,C,K}^s = \{y_{G,C,K,j}^s\}_{j=1}^{Ns} \) of multi-channel signals at \( s \) are found using Eq. (9), where \( C = 1, 2, \ldots, s \):

\[
y_{G,C,K,j}^s = \frac{1}{s} \sum_{i=(j-1)s+C}^{js+C-1} (x_{K,i} - \bar{x}_{K,i})^2 \quad (9)
\]

\[
\bar{x}_{K,i} = \frac{1}{s} \sum_{i=(j-1)s+C}^{js+C-1} x_{K,i}
\]

2. For each \( Y_{G,C,K}^s \), similarity probabilities \( P_{G,C,m}^s(r) \) and \( P_{G,C,m+1}^s(r) \) are calculated according to the MMSE steps (2) to (5).

3. The final MGRCMSE values can be computed by using refined entropy operation:

\[
\text{MGRCMSE}(X_K, s, m, r, N) = - \ln \left[ \frac{\sum_{C=1}^{s} P_{G,C,m+1}^s(r)}{\sum_{C=1}^{s} P_{G,C,m}^s(r)} \right] \quad (10)
\]

2.4 Simulation analysis

The MGRCMSE parameter settings are determined by carrying out noise experiments, allowing the authors to compare its performance and effectiveness to the MMSE, MRCMSE, and MGMSE.

2.4.1 MGRCMSE parameter selection

In the MGRCMSE, four parameters including the data length \( N \), embedding dimension \( m \), similarity tolerance \( r \), and scale factor \( s \) must be set manually. Moreover, to study the influence these parameters have on the MGRCMSE, three-channel white noise (wn) and three-channel pink noise (pn) were utilized.

1. *Influence of the scale factor \( s \) on MGRCMSE.* There is no standard for adjusting the parameter \( s \); therefore, in this paper, the authors set its value to \( s = 25 \).

2. *Influence of the data length \( N \) on MGRCMSE.* The MGRCMSE was used to analyze three-channel noise signals with different lengths: \( N = \{2000, 3000, 4000, 5000\} \). The entropy curve results for
4. The influence of the similarity tolerance $r$ on MGRCMSE. Four different $r$ values (0.1sd, 0.15sd, 0.2sd, and 0.25sd) were used to find the optimal one; the MGRCMSE was applied to analyze the above-mentioned noise signals. Figure 4 illustrates the entropy curves for $N = 3000$, $m = 2$, and $s = 25$.

Figure 4 shows that for the same noise signal, the MGRCMSE entropy curve decreases with the increase in $r$. This can be explained by the fact that, when the value of $r$ is low, the number of matching templates increases, increasing the entropy values. Moreover, when $r$ has a small value (e.g., $r = 0.1sd$), the MGRCMSE entropy curves of both noise signals fluctuate slightly. On the other hand, the number of matching templates decreases when $r$ is a large value, reducing the entropy values. When $r$ is large (e.g., $r = 0.25sd$), the MGRCMSE entropy curves of both noise signals are relatively close in the large scales. When $r$ is 0.15sd or 0.2sd, the MGRCMSE entropy curves of both signal groups are relatively smooth and the noise signals can be distinguished on large scales. Therefore, $r = 0.15sd$ was used in this paper.

To conclude, within the paper, the MGRCMSE parameters are set to $N = 3000$, $m = 2$, $r = 0.15sd$, and $s = 25$.

2.4.2 Comparison of the various multi-channel signal analysis algorithms

In this section, the MGRCMSE algorithm was compared to the MMSE, MRCMSE, and MGMSE algorithms, aiming to assess its effectiveness. Four three-channel signals (as shown in Table 1) were analyzed using each of the algorithms, and the analysis results are shown in Fig. 5. It should be noted that all the algorithm parameter values were consistent with the MGRCMSE parameters.

Several conclusions are drawn from Fig. 5. Firstly, the entropy curves obtained by the MGRCMSE (or MRCMSE) for each group of signals have displayed lower fluctuation compared to the MGMSE (or MMSE). This is since MGRCMSE (or MRCMSE) adopts the multivariate refined composite coarse-grained structure, which allows it to comprehensively consider multiple sequences at $s$. On the other hand, the MGMSE (or MMSE) only considers the single coarse-grained sequence, ignoring the other important

$m = 2$, $r = 0.15sd$ (where $sd$ is the time series standard deviation), and $s = 25$ are plotted in Fig. 2.

According to Fig. 2, the fusion entropy values of the three-channel white noise are higher than those of the three-channel pink noise for the majority of $N$ values. The reason for this phenomenon is that, compared to pink noise, white noise has stronger randomness and weaker self-similarity; therefore, its entropy values are higher. Furthermore, the three-channel white noise (or three-channel pink noise) entropy curves are similar to each other regardless of $N$, meaning that they have a limited effect on the MGRCMSE. Thus, the noise signal length was set to $N = 3000$.

3. The influence of the embedding dimension $m$ on MGRCMSE. By selecting four different $m$ values as $m = \{1, 2, 3, 4\}$, the MGRCMSE was used to analyze two above-mentioned noise signals. The resulting entropy curves obtained using $N = 3000$, $r = 0.15sd$, and $s = 25$ are shown in Fig. 3.

Figure 3 shows that, primarily, both the white and pink noise entropy values increase with the increase in $m$. However, the overall trend practically remains the same. Secondly, when $m$ is a small value (for example, $m = 1$), the white and pink noise entropy curves have similar values on large scales. On the other hand, when $m$ is a large value (i.e., $m = 3$ or 4), the entropy curves fluctuate noticeably. Finally, when $m = 2$, the MGRCMSE entropy curves of both noise signals are relatively smooth, making them more distinguishable on the larger scales. Therefore, the value $m = 2$ was used in this paper.
MCSs information. Thus, the MGRCMSE (or MRCMSE)-obtained entropy values were more stable, further confirming the advantages of the multivariate refined composite coarse-grained structure.

Next, for the MRMCSE (or MMSE), the entropy values of the four group signals at most scales are ordered as $E_{3pn} > E_{1wn2pm} > E_{2wn1pm} > E_{3wn}$, while in the MGRCMSE (or MGMSE), it is
In practice, the signal irregularity values are ordered as $I_{3wn} > I_{2wn1pn} > I_{1wn2pn} > I_{3pn}$. Hence, the MGRCMSE (or MGMSE) algorithm analysis result with multivariate generalized coarse-grained structure is closer to the actual situation.

Lastly, compared to the MMSE, MGMSE, and MRCMSE, the proposed MGRCMSE can not only obtain smooth and easily distinguishable entropy curves but also provide the analysis results for the actual composite situation. Thus, significant improvements were identified when using MGRCMSE when measuring the complexity of multi-channel signals.

### 3 The proposed intelligent rotating machinery fault diagnosis method

In this section, a novel intelligent fault diagnosis method for rotating machinery was proposed. Following the proposal, related algorithms such as SIM and HHO-SVM were described in detail.

#### 3.1 The proposed fault diagnosis scheme

The authors aimed to provide a precise and reliable rotating machinery fault diagnosis process under different working conditions (i.e., a robust process). Thus, an intelligent multi-channel fault diagnosis method for the rotating machinery was proposed, primarily to enable the multi-channel signal fault diagnosis by using MGRCMSE, SIM, and HHO-

### Table 1 Four groups of three-channel noise signals

| Abbreviation | Description                                      |
|--------------|--------------------------------------------------|
| 3 wn         | Three-channel white noise                        |
| 3 pn         | Three-channel pink noise                         |
| 1 wn 2 pn    | One-channel white noise and two-channel pink noise|
| 2 wn 1 pn    | Two-channel white noise and one-channel pink noise|

![Fig. 5](image-url)
The flowchart of the proposed method is shown in Fig. 6.

The specific steps needed to carry out the process are summarized as follows:

1. **Multi-channel signal acquisition.** The multi-channel vibration acceleration signals of rotating machinery were collected through multiple sensors and in various working states. Then, the collected samples are randomly divided into training samples (10%) and testing samples (90%).

2. **Fusion entropy feature extraction.** The proposed MGRCMSE method was used to extract fusion entropy features from multi-channel signals, allowing the extraction of a nonlinear and high-dimensional fusion feature set.

3. **Dimensionality reduction.** The SIM algorithm was introduced to map the MGRCMSE set to the low-dimensional space, obtaining a low-dimensional and sensitive feature set.

4. **Pattern recognition.** The training feature set and testing feature set were normalized to the interval [0, 1], respectively. The training set was used as input to the HHO-SVM, allowing it to search for the optimal parameters ($c_{best}$, $b_{best}$), allowing the authors to obtain the optimized SVM model. Finally, the testing set was input to the trained SVM classifier to diagnose faults.

### 3.2 Supervised Isomap manifold learning method

In the next step, the SIM algorithm was utilized to reduce the MGRCMSE feature set dimensionality, aiming to construct the sensitive, low-dimensional, and high-distinguish feature set. For a given data set $U = \{u_i\}_{i=1}^N$, the key SIM procedures are as follows:
1. First, the supervised distance matrix \( D_i = \{d_i(u_i, u_j)\} \) is calculated:

\[
d_i(u_i, u_j) = \begin{cases} 
\sqrt{1 - \exp\left(-\frac{d^2(u_i, u_j)}{\sigma}\right)}, & L_u = L_j \\
\sqrt{\exp\left(-\frac{d^2(u_i, u_j)}{\sigma}\right)} - \eta, & L_u \neq L_j
\end{cases}
\]

where \( d(u_i, u_j) \) is the Euclidean distance between \( u_i \) and \( u_j \), \( L_u \) is the \( u_i \) category, \( \sigma \) is the average value of \( d(u_i, u_j) \), and \( \eta \) is the weighted coefficient.

2. The neighborhood graph is constructed using the \( k \)-nearest neighbor method. If two sample points (e.g., \( u_i \) and \( u_j \)) are adjacent to each other, there is a side connection with a side length \( d_i(u_i, u_j) \); otherwise, there is no side connection.

3. The shortest path between any two samples can be calculated using the Dijkstra algorithm; the path is approximately regarded as the geodesic distance between these samples.

4. The geodesic distance matrix is mapped to the low-dimensional space by using the multi-dimensional scaling algorithm [28], allowing us to obtain the final dimensionality reduction result.

3.3 Harris hawks optimization-based support vector machine

The SVM, based on statistical learning theory, is the most commonly used classifier. It not only deal with small samples and nonlinear classification problems but also has good generalization ability. However, its performance is closely related to two important parameters—\( c \) and \( g \). Thus, to improve the SVM performance, this section introduces the HHO-SVM classifier obtained after using the novel HHO algorithm to find the optimal values of SVM parameters. The HHO-SVM is a novel classifier (flowchart shown in Fig. 7) utilizing the Harris hawks optimization algorithm (for more details, see [29]) to optimize the key SVM parameters. The specific steps are shown next.

1. **Preprocess the data** Equation (12) is applied to normalize the input data set between [0, 1].

\[
v' = (v - v_{\text{min}})/(v_{\text{max}} - v_{\text{min}})
\]

where \( v \) and \( v' \) are the values before and after normalization, respectively. Similarly, \( v_{\text{max}} \) and \( v_{\text{min}} \) are the maximum and minimum values, respectively.

2. **Initialize the parameters**. The initial HHO population size \( (N_p) \) is set to 20, and the maximum number of iterations \( (T) \) is 100. For the optimization problem at hand, the main purpose is to find the optimal SVM parameter values—\( c \) and \( g \). The position of each Harris hawk is defined using \( H_i = (c, g) \), with the lower and upper position boundaries set as \( LB = [0.001, 0.001] \) and \( UB = [100, 100] \), respectively.

3. **Calculate the fitness values**. To better evaluate the quality of each Harris hawk, the average training sample error classification rate after a threefold crossover is taken as the fitness value. In other words, the training samples are divided into three subsets; each subset is regarded as a verification set, while the two remaining sets are used as training sets. Finally, the average error classification rate of the three groups is taken as the fitness value.

4. **Search the Prey location**. The best Harris hawk position, which corresponds to the best fitness value, is considered as the prey location \( (H_{\text{rabbit}}) \) under the current iteration \( t \).

5. **Update the prey escaping energy and jump strength**. The escaping energy \( E \) and the jump strength \( J \) of the prey can be updated, using Eq. (13):

\[
E = 2E_0(1 - t/T)
\]

\[
J = 2(1 - r_1)
\]

\[
E_0 = 2r_2 - 1
\]

where \( r_1 \) and \( r_2 \) are random numbers between (0, 1).

6. **Update the Harris hawk locations**. The position of each Harris hawk can be updated by using Eqs. (14) to (18). If the fitness value of the new position is lower than that of the previous position, the previous position is replaced with the new one; otherwise, the previous position is retained.

1. **Survey phase** (i.e., \( |E| \geq 1 \)). The position of each Harris hawks is updated by:
where $H(t)$ and $H(t+1)$ are the Harris hawk positions at $t$ and $t+1$ iteration, respectively. $H_{rabbit}(t)$ is the prey location, $H_{rand}(t)$ is the random Harris hawk position in the current iteration, and $H_m = \frac{1}{N_p} \sum_{i=1}^{N_p} H_i(t)$ is the average position of all Harris hawks positions in the current iteration, while $r_3, r_5, r_6$ and $q$ are the random numbers between 0 and 1.

2. Development phase (i.e., $0 \leq |E| < 1$). The variable $r_7$ is defined as the probability of the prey escaping and is a random number between 0 and 1. For example, $r_7 < 0.5$ indicates that the prey has succeeded in escaping while $r_7 \geq 0.5$ indicates that it has failed. This phase can be further into four possible situations:

- **Soft besiege** (i.e., $r_7 \geq 0.5, |E| \geq 0.5$)—the position of each Harris hawk is updated by:

  $$H(t+1) = \Delta H(t) - E[\Delta H_{rabbit}(t) - H(t)]$$

  $$\Delta H(t) = H_{rabbit}(t) - H(t)$$

- **Hard besiege** (i.e., $r_7 \geq 0.5, 0 \leq |E| < 0.5$)—the position of each Harris hawk is updated by:

  $$H(t+1) = H_{rabbit}(t) - E[\Delta H(t)]$$

- **Soft besiege with progressive rapid dives** (i.e., $r_7 < 0.5, |E| \geq 0.5$)—the position of each Harris hawk is updated by:

  $$H(t+1) = H_{rabbit}(t) - E[\Delta H(t)]$$

  $$\Delta H(t) = H_{rabbit}(t) - H(t)$$

  $$\Delta H(t) = H_{rabbit}(t) - H(t)$$

- **Hard besiege with progressive rapid dives** (i.e., $r_7 < 0.5, |E| < 0.5$)—the position of each Harris hawk is updated using:

  $$H(t+1) = \begin{cases} 
  Y_u = H_{rabbit}(t) - E[\Delta H(t)] & \text{if } F(Y_u) < F(H(t)) \\
  Y_d = Y_u + S \times LF(x) & \text{if } F(Y_d) < F(H(t)) \\
  \end{cases}$$

  $$LF(x) = 0.01 \times \mu \times \theta$$

  $$\theta = \left( \left[ \frac{\Gamma(1 + \beta) \times \sin \left( \frac{\pi \beta}{2} \right) \times \beta \times 2^{1/\beta} \right] \right)^{1/\beta}$$

  where $D = 2$ is the dimensional position, $S$ is a random vector of size $1 \times D$, $LF(x)$ is the Levy flight function, and $\mu, \gamma$ are the random number between 0 and 1. Finally, $\beta = 1.5$ and $\Gamma(\cdot)$ is a Gamma function.

- **Hard besiege with progressive rapid dives** (i.e., $r_7 < 0.5, |E| < 0.5$)—the position of each Harris hawk is updated using:

  $$H(t+1) = \begin{cases} 
  Y_u = H_{rabbit}(t) - E[\Delta H(t)] & \text{if } F(Y_u) < F(H(t)) \\
  Y_d = Y_u + S \times LF(x) & \text{if } F(Y_d) < F(H(t)) \\
  \end{cases}$$
7. Determine whether the stop condition is met. When the simulation reaches the previously selected maximum number of iterations, the cycle is terminated. The final prey position \((c_{\text{best}}, g_{\text{best}})\) is used as the output. Otherwise, step (2) is repeated until the stop criterion is met.

8. Construct the optimal SVM training model. The final prey position is implemented in the SVM prediction model.

9. Identify the testing set categories. The testing samples are input into the trained SVM prediction model for classification with the recognition results as outputs.

4 Experimental verification

The experiments were carried out using two case studies as examples: the first being concerned with bearings and the second with gears. The aim was to evaluate the performance and verify the effectiveness of the proposed fault diagnosis method. Moreover, several comparative experiments were carried out to evaluate the proposed method and verify its advantages. These include both the comparison with other multi-channel feature extraction methods and the comparison with several single-channel feature extraction methods.

4.1 Experimental rig

The experimental data were collected from the drivetrain dynamics simulator (Fig. 8). The experimental platform consists of the variable speed drive motor, programmable motor controller, planetary gearbox, parallel shaft gearbox, brake, and programmable magnetic controller. The planetary gearbox parameters are given in Table 2.

4.2 Case study 1: Rolling bearing fault diagnosis

4.2.1 Rolling bearing data set

Four bearing operation states were considered in this experiment: normal, inner ring fault, outer ring fault, and ball fault, with each being outlined in Fig. 9. The working experimental rig conditions were selected as follows: condition 1—the driving speed of 25 Hz and no load; and condition 2—the driving speed of 35 Hz and the 2 V load. The acceleration sensors were used to collect the vibration signals in the X, Y, and Z directions, each obtained with the sampling frequency of 3000 Hz (please see Table 3). Finally, the three-channel time-domain bearing waveforms under various operation states are shown in Figs. 10 and 11.

4.2.2 Experimental results and analysis on condition 1

According to the proposed fault diagnosis process, the MGRCMSE method was applied to extract the fusion features from three-channel bearing signals operating under condition 1. The results for 800 bearing samples analyzed under four different states are plotted in Fig. 12.

Several phenomena are found in Fig. 12. Firstly, the mean entropy values of the normal bearing signals are higher when compared to the fault bearing signals, regardless of the scale. Such behavior is caused by the self-similarity of normal signals, which is weaker than that of fault signals. Therefore, the normal state entropy values are higher than their fault state counterparts. Hence, the MGRCMSE algorithm can effectively monitor the occurrence of bearing failure. Secondly, the entropy values of the four operation states at most scales are ordered as follows: \(E_{\text{normal}} > E_{\text{BF}} > E_{\text{IRF}} > E_{\text{ORF}}\), which is consistent with the realistic case. Both phenomena confirm that the MGRCMSE algorithm can identify the faults.

However, Fig. 12 shows that the IRF and BF entropy mean curves are rather similar. Moreover, the extracted feature set is high-dimensional and redundant; if directly input to the classifier as such, both the recognition time and error will be increased. For this reason, according to the proposed fault diagnosis method, a supervised manifold learning algorithm (SIM) was adopted to generate the corresponding low-dimensional set. The results are shown in Fig. 13. It should be pointed out that based on several experiments, the optimal SIM parameters are \(d = 3, k = 92, \) and \(\eta = 0.4\).

In the 3D visualization results (see Fig. 13), it was shown that SIM could entirely separate the four types of samples. Additionally, it is evident that samples are grouped based on their state. This proves that the SIM algorithm can effectively obtain a low-dimensional and sensitive set, which is easily distinguishable.
When classifying the set, the above-presented fusion entropy manifold (i.e., MGRCMSE + SIM) feature set was used as input for the HHO-SVM classifier. The HHO algorithm was first used to find the optimal SVM parameters (i.e., \( c_{\text{best}} = 79.86 \), \( g_{\text{best}} = 7.62 \)), allowing us to establish the optimal SVM prediction model. The training stage fitness curve is shown in Fig. 14a. Testing samples were input to the previously trained SVM classifier for pattern recognition; output results and the confusion matrix are shown in Fig. 14b and c, respectively.

As shown in Fig. 14, 720 testing samples were correctly identified with an average recognition rate of 100%. More importantly, it was proved that the proposed fault diagnosis method could effectively and accurately diagnose bearings in different states.

| Component   | Number of gear teeth | First stage | Second stage |
|-------------|----------------------|-------------|--------------|
| Gear ring   | 100                  | 100         |              |
| Planet gear | 40 (the number of planetary gears is 3) | 36 (the number of planetary gears is 4) |
| Sun gear    | 20                   | 28          |              |

Fig. 8 Experimental test platform

Table 2 Main parameters of planetary gearbox

Fig. 9 Bearing faults: a inner ring fault, b outer ring fault, and c ball fault
Table 3  Description of rolling bearings under four working states

| State                | Abbreviation | Description                  | Class | Channel | Data points | Number of samples | Number of training samples | Number of testing samples |
|----------------------|--------------|------------------------------|-------|---------|-------------|--------------------|---------------------------|--------------------------|
| Normal               | NOR          | No fault                     | 1     | X       | 3000        | 200                | 20                        | 180                      |
|                      |              |                              |       | Y       | 3000        |                    |                           |                          |
|                      |              |                              |       | Z       | 3000        |                    |                           |                          |
| Inner ring fault     | IRF          | A fault occurs in the inner ring | 2     | X       | 3000        | 200                | 20                        | 180                      |
|                      |              |                              |       | Y       | 3000        |                    |                           |                          |
|                      |              |                              |       | Z       | 3000        |                    |                           |                          |
| Outer ring fault     | ORF          | A fault occurs in the outer ring | 3     | X       | 3000        | 200                | 20                        | 180                      |
|                      |              |                              |       | Y       | 3000        |                    |                           |                          |
|                      |              |                              |       | Z       | 3000        |                    |                           |                          |
| Ball fault           | BF           | A fault occurs in the ball    | 4     | X       | 3000        | 200                | 20                        | 180                      |
|                      |              |                              |       | Y       | 3000        |                    |                           |                          |
|                      |              |                              |       | Z       | 3000        |                    |                           |                          |

Fig. 10  Time-domain waveforms of rolling bearing signals at condition 1: a NOR, b IRF, c ORF, and d BF

Fig. 11  Time-domain waveforms of rolling bearing signals at condition 2: a NOR, b IRF, c ORF, and d BF
To investigate the effect of operating conditions on the proposed fault diagnosis method, it was used to analyze the bearing data under operating condition 2. The MGRCMSE method was first applied to extract the fusion features from three-channel bearing signals. The results for a total of 800 analyzed bearing samples under four different states are plotted in Fig. 15.

Figure 15 shows that when the bearing conditions change, the normal state entropy curve is higher than that of the fault states (at larger scale numbers). Moreover, the entropy values of four analyzed fault states at most scales are ordered as

4.2.3 Experimental results and analysis on condition 2

The MGRCMSE method was first applied to extract the fusion features from three-channel bearing signals. The results for a total of 800 analyzed bearing samples under four different states are plotted in Fig. 15.

Figure 15 shows that when the bearing conditions change, the normal state entropy curve is higher than that of the fault states (at larger scale numbers). Moreover, the entropy values of four analyzed fault states at most scales are ordered as
which is consistent with the realistic case. The underlying reasons for such behavior are identical to those given in Sect. 4.2.2. This experiment further demonstrates the advantages of the proposed fault feature extraction method.

Subsequently, the SIM algorithm was used to reduce the dimensionality of the collected MGRCMSE feature set, as shown in Fig. 16. It should be pointed out that the optimal SIM parameters were \(d = 3\), \(k = 200\), and \(\eta = 0.35\). As shown in Fig. 16, the SIM method can be used to effectively separate four bearing states and obtain both low-dimensional and sensitive fault features.

Finally, the HHO-SVM classifier was used to identify faults of the obtained MGRCMSE + SIM set. The HHO algorithm was applied to find the optimal SVM parameters (i.e., \(c_{\text{best}} = 33.44\), \(g_{\text{best}} = 2.15\)) necessary for building the optimal SVM prediction model. The training stage fitness curve is shown in Fig. 17a. Testing samples were then inputted to the previously trained SVM classifier for pattern recognition; output results and testing sample confusion matrices are shown in Figs. 17b and c, respectively.

It can be concluded that 720 testing samples were correctly identified with an average recognition rate of 100%. More importantly, it was proved that the proposed fault diagnosis method could, effectively and accurately, diagnose bearings states in various working conditions.

4.3 Case study 2: Gear fault diagnosis

4.3.1 Gear data set

In this experiment, four gear operation states were established: normal, tooth breakage fault, missing tooth fault, and tooth root fault (Fig. 18). The experimental working conditions were as follows: the driving speed of 25 Hz and no load. Three acceleration sensors were used to collect the vibration signals in X, Y, and Z directions under different operation modes, obtained at the sampling frequency of 3000 Hz. The detailed description is shown in Table 4, while the three-channel gear waveforms under each operation mode are plotted in Fig. 19.
Fig. 17 Fault diagnosis results: a fitness curve, b output results, and c confusion matrix (%)

Fig. 18 Gear faults: a tooth root fault, b tooth breakage fault, and c missing tooth fault

Table 4 Description of gears under four working states

| State            | Abbreviation | Description          | Class | Channel | Data points | Number of samples | Number of training samples | Number of testing samples |
|------------------|--------------|----------------------|-------|---------|-------------|--------------------|--------------------------|--------------------------|
| Normal           | NOR          | Healthy              | 1     | X       | 3000        | 200                | 20                       | 180                      |
|                  |              |                      |       | Y       | 3000        |                    |                          |                          |
|                  |              |                      |       | Z       | 3000        |                    |                          |                          |
| Tooth breakage   | TBF          | Signal tooth breakage | 2     | X       | 3000        | 200                | 20                       | 180                      |
| fault            |              |                      |       | Y       | 3000        |                    |                          |                          |
|                  |              |                      |       | Z       | 3000        |                    |                          |                          |
| Missing tooth    | MTF          | Signal tooth missing | 3     | X       | 3000        | 200                | 20                       | 180                      |
| fault            |              |                      |       | Y       | 3000        |                    |                          |                          |
|                  |              |                      |       | Z       | 3000        |                    |                          |                          |
| Tooth root       | TRF          | A fault occurs in the gear tooth root | 4     | X       | 3000        | 200                | 20                       | 180                      |
| fault            |              |                      |       | Y       | 3000        |                    |                          |                          |
|                  |              |                      |       | Z       | 3000        |                    |                          |                          |
4.3.2 Experimental results and analysis

According to the proposed fault diagnosis method, the MGRCMSE algorithm was utilized to extract the fusion features from three-channel gear signals, as shown in Fig. 20.

The following observations are made based on Fig. 20: (1) The mean entropy values of normal gear signals were higher than that of faulty gear signals at most of the scales. The main reason for such behavior was the occurrence of vibration signals in the normal state, which have shown strong randomness, while the vibration signals in faulty states had a prolonged impact. Thus, the entropy values of fault states were smaller than that of the normal state, indicating that the MGRCMSE can monitor the occurrence of gear failure. (2) The mean entropy curves of the gear signals under four states can be separated from each other, showing that the proposed method can effectively distinguish between the fault states.

The SIM was used to reduce the extracted MGRCMSE feature set dimensionality. (The result is plotted in Fig. 21.) It should be pointed out that the optimal SIM parameters were obtained experimentally and are as follows: $d = 3$, $k = 193$, and $\eta = 0.38$.

The 3D visualization results (shown in Fig. 21) show that SIM thoroughly separates the four-state samples, completely avoiding sample overlap. Thus, the SIM effectiveness in the dimensionality reduction of the gear feature set was considered as verified.

Furthermore, the obtained fusion entropy manifold (i.e., MGRCMSE + SIM) feature set was used as input for the classification using HHO-SVM. The HHO algorithm was firstly employed to find the optimal SVM parameters ($c_{\text{best}} = 56.99$, $g_{\text{best}} = 24.84$), allowing us to establish the corresponding optimal SVM prediction model. The training stage fitness curve is shown in Fig. 22a, while the testing sample output results and the confusion matrix are shown in Fig. 22b, c.
According to Fig. 23a–d, the mean MGMSE and MGRCMSE entropy curves for each bearing state are relatively similar. However, it is noted that the mean MGMSE entropy curves have slight fluctuations in large scales, especially for bearing without faults. Next, Fig. 23e–h shows that the entropy values obtained using MGRCMSE have low standard deviations on most scales. Additionally, several random entropy curves were drawn for each bearing state signal, using both algorithms, as shown in Fig. 23i–l. When compared with the MGRCMSE, the MGMSE entropy curves have larger fluctuations. These phenomena prove that the proposed MGRCMSE algorithm can obtain more stable entropy results when compared to the MGMSE.

Furthermore, the feature sets obtained using the above-mentioned algorithms were input into the HHO-SVM for recognition (for results, please see Fig. 24).

As illustrated in Fig. 24, a total of 126 MGRCMSE testing samples were misclassified, including four NOR samples, 42 IRF samples, 13 ORF samples, and 67 BF samples. For a comparison, when using MGMSE, 148 testing samples were not correctly classified, including nine NOR samples, 61 IRF samples, six ORF samples, and 72 BF samples. Thus, compared to the MGRCMSE (the average recognition rate of 82.50%), the average MGMSE recognition rate was 3.06% lower (i.e., 79.44%). Thus, it is verified that the MGRCMSE performance in feature extraction is an improvement compared to MGMSE.

4.4 Performance comparison

The advantages of the proposed MGRCMSE method were then further explored using the bearing fault diagnosis as an example; two comparative studies were carried out.

1. Comparison with other multi-channel feature extraction algorithm. In this section, the MGRCMSE algorithm performance was compared to the performance of the MGMSE algorithm; the results are shown in Fig. 23. The MGMSE algorithm parameters in this study are listed in Table 5.

2. Comparison with single-channel feature extraction algorithms. In this section, the MGRCMSE algorithm performance was compared to GMSE, GRCMSE, generalized composite multi-scale
fuzzy entropy (GCMFE), and generalized composite multi-scale permutation entropy (GCMPE). These single-channel feature extraction methods were used to extract features from the X, Y, and Z channel signals. The extracted results were then submitted to the HHO-SVM classifier for fault identification and the results are shown in Fig. 25. The algorithm parameters used to carry out the study are given in Table 6.

As shown in Fig. 25, the diagnosis accuracy of the same single-channel feature extraction algorithms (i.e., GMSE, GRCMSE, GCMFE, and GCMPE) varied significantly depending on the channel signals. For example, the GRCMSE algorithm had the highest average recognition accuracy for channel X signals (i.e., 81.25%), while it was the lowest for channel Y signals (i.e., 78.19%). The total difference was 3.06%. Moreover, the proposed MGRCMSE method had the

---

**Table 5** Parameter settings for various feature extraction methods

| Reference               | Algorithms | Type               | Parameter settings                      |
|-------------------------|------------|--------------------|----------------------------------------|
| Yin et al. [26]         | MGMSE      | Multi-channel method | \(m = 2, r = 0.15sd, N = 3000, \) and \(s = 25\) |
| Wang et al. (proposed)  | MGRCMSE    | Multi-channel method | \(m = 2, r = 0.15sd, N = 3000, \) and \(s = 25\) |
highest recognition accuracy for bearing fault signals. Thus, it was shown that MGRCMSE can both fully exploit the fusion features of all the channels and explore the fault information more thoroughly compared to existing single-channel feature extraction methods. Lastly, the proposed MGRCMSE algorithm can avoid the channel signal selection effects present in single-channel feature extraction algorithms.

3. Comparison with similar fault diagnosis methods. To validate the potential application of the proposed fault diagnosis method, it was compared to other previously published state-of-the-art methods [33–38], as shown in Table 7. The

![Figure 24](image1)
![Figure 25](image2)

**Fig. 24** Fault diagnosis results of two feature extraction sets using the HHO-SVM: a MGRCMSE fitness curve, b MGRCMSE output results, c MGRCMSE confusion matrix (%), d MGMSE fitness curve, e MGMSE output results, and f MGMSE confusion matrix (%)

**Fig. 25** Fault diagnosis accuracy of various feature extraction algorithms
The proposed MGRCMSE algorithm can effectively extract the fusion entropy features for both bearings and gears.

2. Compared to the existing multi-channel feature extraction algorithm (i.e., MGMSE) and single-channel feature extraction algorithms (i.e., GMSE, GRCMSE, GCMFE, GCMPE), the MGRCMSE displayed the highest fault diagnostic accuracy.

3. The proposed intelligent multi-channel fault diagnosis method, based on MGRCMSE, can accurately diagnose both the bearing and gear faults.
When the rotating machinery has faults, the corresponding vibration signals will show regular transients. The proposed MGRCMSE method was used for rotating machinery fault diagnosis by exploring the complexity and regularity of multi-channel vibration signals. However, the MGRCMSE method parameters rely on manual or empirical settings. Thus, in the future, the authors aim to focus on mitigating this problem, and further apply it to diagnose similar rotating machinery faults.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China (Grant No. 51775114, 51875105, 51275092); the Fujian Provincial Industrial Robot Basic Components Technology Research and Development Center (Grant No. 2014H21010011); and the Natural Science Foundation of Anhui Province (Grant No. 1808085ME152).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Data availability The data sets generated during and analyzed during the current study are available from the corresponding authors on reasonable request.

References

1. Prabith, K., Praveen Krishna, I.R.: The numerical modeling of rotor-stator rubbing in rotating machinery: a comprehensive review. Nonlinear Dyn. 101, 1317–1363 (2020). https://doi.org/10.1007/s11071-020-05832-y
2. Zheng, J., Pan, H.: Use of generalized refined composite multiscale fractional dispersion entropy to diagnose the faults of rolling bearing. Nonlinear Dyn. 101, 1417–1440 (2020) https://doi.org/10.1007/s11071-020-05821-4
3. Li, X., Li, J., Zhao, C., Qu, Y., He, D.: Gear pitting fault diagnosis with mixed operating conditions based on adaptive 1D separable convolution with residual connection. Mech. Syst. Signal Process. 142, 106740 (2020). https://doi.org/10.1016/j.ymssp.2020.106740
4. Pang, S., Yang, X., Zhang, X., Lin, X.: Fault diagnosis of rotating machinery with ensemble kernel extreme learning machine based on fused multi-domain features. ISA Trans. 98, 320–337 (2020). https://doi.org/10.1016/j.isatra.2019.08.053
5. Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., Nandi, A.K.: Applications of machine learning to machine fault diagnosis: A review and roadmap. Mech. Syst. Signal Process. 138, 106587 (2020). https://doi.org/10.1016/j.ymssp.2019.106587
6. Xue, Y., Dou, D., Yang, J.: Multi-fault diagnosis of rotating machinery based on deep convolution neural network and support vector machine. Measurement 156, 107571 (2020). https://doi.org/10.1016/j.measurement.2020.107571
7. Zhao, B., Zhang, X., Li, H., Yang, Z.: Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions. Knowl. Based Syst. 199, 105971 (2020). https://doi.org/10.1016/j.knosys.2020.105971
8. Zhang, Z., Li, S., Lu, J., Wang, J., Jiang, X.: A novel intelligent fault diagnosis method based on fast intrinsic component filtering and pseudo-normalization. Mech. Syst. Signal Process. 145, 106923 (2020). https://doi.org/10.1016/j.ymssp.2020.106923
9. Li, Y., Wang, S., Deng, Z.: Intelligent fault identification of rotary machinery using refined composite multi-scale Lempel–Ziv complexity. J. Manuf. Syst. (2020) (in Press) https://doi.org/10.1016/j.jmsy.2020.05.004
10. Wu, Y., Jiang, P., Ding, C., Feng, F., Chen, T.: Intelligent fault diagnosis of rotating machinery based on one-dimensional convolutional neural network. Comput. Ind. 108, 53–61 (2019). https://doi.org/10.1016/j.compind.2018.12.001
11. Zhao, X., Jia, M., Ding, P., Yang, C., She, D., Liu, Z.: Intelligent fault diagnosis of multi-channel motor-rotor system based on multi-manifold deep extreme learning machine. IEEE/ASME Trans. Mechatron. 25(5), 2177–2187 (2020). https://doi.org/10.1109/TMECH.2020.3004589
12. Shao, S., Yan, R., Lu, Y., Wang, P., Gao, R.X.: DCNN-based multi-signal induction motor fault diagnosis. IEEE Trans. Instrum. Meas. 69(6), 2658–2669 (2020). https://doi.org/10.1109/TIM.2019.2925247
13. Wu, J., Jiang, B., Chen, H., Liu, J.: Sensors information fusion system with fault detection based on multi-manifold regularization neighborhood preserving embedding. Sensors 19(6), 1440 (2019). https://doi.org/10.3390/s19061440
14. Wang, Z., Yao, L., Cai, Y., Zhang, J.: Mahalanobis semi-supervised mapping and beetle antennae search based support vector machine for wind turbine rolling bearings fault diagnosis. Renew. Energy 155, 1312–1327 (2020). https://doi.org/10.1016/j.renene.2020.04.041
15. Teng, Y., Shang, P., He, J.: Multiscale fractional-order approximate entropy analysis of financial time series based on the cumulative distribution matrix. Nonlinear Dyn. 97, 1067–1085 (2019). https://doi.org/10.1007/s11071-019-05033-2
16. Liu, H., Han, M.: A fault diagnosis method based on local mean decomposition and multi-scale entropy for roller bearings. Mech. Mach. Theory 75, 67–78 (2014). https://doi.org/10.1016/j.mechmachtheory.2014.01.011
17. Landauskas, M., Cao, M., Ragulskis, M.: Permutation entropy-based 2D feature extraction for bearing fault diagnosis. Nonlinear Dyn. 102, 1717–1731 (2020). https://doi.org/10.1007/s11071-020-06014-6
18. Zheng, J., Jiang, Z., Pan, H.: Sigmoid-based refined composite multi-scale fuzzy entropy and t-SNE based fault diagnosis approach for rolling bearing. Measurement 129, 332–342 (2018). https://doi.org/10.1016/j.measurement.2018.07.045
19. Yan, X., Jia, M.: Intelligent fault diagnosis of rotating machinery using improved multiscale dispersion entropy and mRMR feature selection. Knowl. Based Syst. 163,
20. Wang, C.: A sample entropy inspired affinity propagation method for bearing fault signal classification. Digital Signal Process. 102, 102740 (2020). https://doi.org/10.1016/j.dsp.2020.102740
21. Gao, Q., Liu, W., Tang, B., Li, G.: A novel wind turbine fault diagnosis method based on integral extension load mean decomposition multi-scale entropy and least squares support vector machine. Renewable Energy 116, 169–175 (2018). https://doi.org/10.1016/j.renene.2017.09.061
22. Dai, J., Zheng, J., Pan, H., Pan, Z.: Rolling bearing fault diagnosis method based on composite multi-scale entropy and Laplacian SVM. China Mech. Eng. 28(11), 1339–1346 (2017). https://doi.org/10.3969/j.issn.1004-132X.2017.11.014
23. Wang, Z., Yao, L., Cai, Y.: Rolling bearing fault diagnosis using generalized refined composite multi-scale sample entropy and optimized support vector machine. Measurement 156, 107574 (2020). https://doi.org/10.1016/j.measurement.2020.107574
24. Lu, Y., Wang, J.: Multi-variate multi-scale entropy of financial markets. Commun. Nonlinear Sci. Numer. Simul. 52, 77–90 (2017). https://doi.org/10.1016/j.cnsns.2017.04.028
25. Humeau-Heurtier, A.: Multi-variate refined composite multi-scale entropy analysis. Phys. Lett. A 380(16), 1426–1431 (2016). https://doi.org/10.1016/j.physleta.2016.02.029
26. Yin, Y., Wang, X., Li, Q., Shang, P.: Generalized multi-variate multi-scale sample entropy for detecting the complexity in complex systems. Phys. A 545, 123814 (2020). https://doi.org/10.1016/j.physa.2019.123814
27. Houssein, E.H., Hosney, M.E., Oliva, D., Mohamed, W.M., Hassaballah, M.: A novel hybrid Harris hawks optimization and support vector machines for drug design and discovery. Comput. Chem. Eng. 133, 106656 (2020). https://doi.org/10.1016/j.compchemeng.2019.106656
28. Machado, J.T., Lopes, A.M.: Multidimensional scaling and visualization of patterns in prime numbers. Commun. Nonlinear Sci. Numer. Simul. 83, 105128 (2020). https://doi.org/10.1016/j.cnsns.2019.105128
29. Heidari, A.A., Mirjalili, S., Faris, H., Aljaraah, I., Mafarja, M., Chen, H.: Harris hawks optimization: Algorithm and applications. Futur. Gener. Comput. Syst. 97, 849–872 (2019). https://doi.org/10.1016/j.future.2019.02.028
30. Costa, M. D., Goldberger, A. L.: Generalized multiscale entropy analysis: Application to quantifying the complex volatility of human heartbeat time series. Entropy 17(3), 1197–1203 (2015) https://doi.org/10.3390/e17031197
31. Wei, Y., Li, Y., Xu, M., Huang, W.: Intelligent fault diagnosis of rotating machinery using ICD and generalized composite multi-scale fuzzy entropy. IEEE Access 7, 38983–38995 (2019). https://doi.org/10.1109/ACCESS.2018.2876759
32. Zheng, J., Pan, H., Yang, S., Cheng, J.: Generalized composite multiscale permutation entropy and Laplacian score based rolling bearing fault diagnosis. Mech. Syst. Signal Process. 99, 229–243 (2018). https://doi.org/10.1016/j.ymssp.2017.06.011
33. Li, J., Xu, Y., Bao, W., Li, Z., Li, L.: Finite-time non-fragile state estimation for discrete neural networks with sensor failures, time-varying delays and randomly occurring sensor nonlinearity. J. Franklin Inst. 356(3), 1566–1589 (2019). https://doi.org/10.1016/j.jfranklin.2018.10.032
34. Li, J., Liu, X., Xu, X., Xu, X.: Disturbance rejection adaptive fault-tolerant constrained consensus for multi-agent systems with failures. IEEE Trans. Circuits Syst. II Express Briefs 67(12), 3302–3306 (2020). https://doi.org/10.1109/TCSII.2020.2986059
35. Zhao, D., Lam, H.K., Li, Y., Ding, S.X., Liu, S.: A novel approach to state and unknown input estimation for takagi–sugeno fuzzy models with applications to fault detection. IEEE Trans. Circuits Syst. I Regul. Pap. 67(6), 2053–2063 (2020). https://doi.org/10.1109/TCSI.2020.2968732
36. Dibaj, A., Ettefagh, M.M., Hassannejad, R., Ehghaghi, M.B.: A hybrid fine-tuned VMD and CNN scheme for untrained compound fault diagnosis of rotating machinery with unequal-severity faults. Expert Syst. Appl. 167, 114094 (2021). https://doi.org/10.1016/j.eswa.2020.114094
37. Zhang, J., Xu, B., Wang, Z., Zhang, J.: An FSK-MBCNN based method for compound fault diagnosis in wind turbine gearboxes. Measurement 172, 108933 (2021). https://doi.org/10.1016/j.measurement.2020.108933
38. Wang, Z., Yao, L., Chen, G., Ding, J.: Modified multiscale weighted permutation entropy and optimized support vector machine method for rolling bearing fault diagnosis with complex signals. ISA Transactions (2021) (In Press) https://doi.org/10.1016/j.isatra.2020.12.054

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.