

Article

Application of Salp Swarm Algorithm and Extended Repository Feature Selection Method in Bearing Fault Diagnosis

Chun-Yao Lee ^{1,*} , Truong-An Le ² , Yung-Chi Chen ³ and Shih-Che Hsu ³

¹ Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei City 106, Taiwan

² Department of Electrical and Electronics Engineering, Thu Dau Mot University, Thu Dau Mot 75000, Vietnam; anlt@tdmu.edu.vn

³ Department of Electrical Engineering, Chung Yuan Christian University, Taoyuan City 320, Taiwan; g11178004@cycu.edu.tw (Y.-C.C.); anthonyh@cycu.edu.tw (S.-C.H.)

* Correspondence: cylee@mail.ntust.edu.tw; Tel.: +886-3-2654827

Abstract: Motor fault diagnosis is an important task in the operational monitoring of electrical machines in manufacturing. This study proposes an effective bearing fault diagnosis model for electrical machinery based on machine learning techniques. The proposed model is a combination of three processes: feature extraction of signals collected from the motor based on multi-resolution analysis, fast Fourier transform, and envelope analysis. Next, redundant or irrelevant features are removed using the feature selection technique. A binary salps swarm algorithm combined with an extended repository is the proposed method to remove unnecessary features. As a result, an optimal feature subset is obtained to improve the performance of the classification model. Finally, two classifiers, *k*-nearest neighbor and support vector machine, are used to classify the fault of the electric motor. There are four input datasets used to evaluate the model performance, and UCI is the benchmark dataset to verify the effectiveness of the proposed feature selection technique. The remaining three datasets include the bearing dataset collected from experiments, with an average classification accuracy of 99.9%, as well as Case Western Reserve University (CWRU) and Machinery Failure Prevention Technology (MFPT), which are public datasets with average classification accuracies of 99.6% and 98.98%, respectively. The experimental results show that this method is more effective in diagnosing bearing faults than other traditional methods and prove its robustness.

Keywords: bearing fault diagnosis; feature extraction; feature selection; machine learning; salp swarm algorithm

MSC: 68W50



Citation: Lee, C.-Y.; Le, T.-A.; Chen, Y.-C.; Hsu, S.-C. Application of Salp Swarm Algorithm and Extended Repository Feature Selection Method in Bearing Fault Diagnosis.

Mathematics **2024**, *12*, 1718. <https://doi.org/10.3390/math12111718>

Academic Editor: Ioannis G. Tsoulos

Received: 24 April 2024

Revised: 20 May 2024

Accepted: 29 May 2024

Published: 31 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rotating machinery is one of the most important types of equipment in industry today. Due to harsh environments, some important parts such as bearings often fail, which has a great impact on the overall operation [1]. Rotating machine failures most commonly occur on rolling bearings [2], and they may not only cause motor failures [3] but there may also even be concerns about work safety. Rolling bearing failures account for 41% of rotating machinery failures [4]. Therefore, bearing fault diagnosis models have attracted more and more attention from researchers. How to solve motor failures and avoid accidents has become a very important issue. In this study, feature selection plays the role of extracting features from the original signal obtained from the induction motor, and then uses feature selection to screen out the best feature subset. Finally, the classification accuracy is obtained by using a classifier.

The purpose of feature extraction is to extract useful signals as features from the original current signal obtained by the induction motor, but the original current signal

contains a lot of noise. The filtered features will change the classification accuracy. Many feature selection methods are widely used, such as multi-resolution analysis (MRA) [5], fast Fourier transform (FFT) [6], envelope analysis (EA) [7], empirical mode analysis (EMD) [8], and Hilbert–Huang transform (HHT) [9]. MRA was proposed by Mallat et al. and is a recursive filtering feature selection method that calculates the wavelet decomposition of the signal based on the finest scale approximation of the signal [10]. FFT is a feature extraction method often used in fault diagnosis. It is mainly used for the analysis and design of discrete signal processing. [11]. EA can obtain the upper envelope and lower envelope of the signal through extreme value sampling and signal reconstruction [12]. This study combines three feature extraction methods: MRA, FFT, and EA. MRA uses the coefficient vector of discrete wavelet transform (DWT) to analyze the original signal and extract time domain features. In addition, it extracts frequency domain features from the signal analyzed by FFT and then uses EA to reorganize the extracted features into feature sets. The feature set generated through feature extraction usually still contains redundant features, especially in high-dimensional feature sets [13], which will affect the final classification and lead to a decrease in accuracy. Therefore, the role of feature selection in the next stage will be crucial.

There are three common approaches in feature selection methods. The filter approach is a method of calculating distances and error rates between features to score or rank them. This approach has the advantage of calculating and capturing important characteristics of the feature. Commonly used methods are Relief [14], minimum redundancy and maximum relevance (mRMR) [15], and correlation-based feature selection (CFS) [16]. The wrapper approach typically uses optimization algorithms to find the optimal feature subset by training a new model for each subset, so they are computationally intensive. Several swarm algorithms belonging to the wrapper approach have been applied in feature selection, such as the salp swarm algorithm (SSA) [17], particle swarm optimization (PSO) [18], gray wolf optimizer (GWO) [19], whale optimization algorithm (WOA) [20], and genetic algorithm (GA) [21]. Each method has its distinct characteristics. Among them, SSA was proposed by Mirjalili et al. The main axis is the behavior of salps looking for food, which pumps water into the body as the driving force for propulsion [17]. Although SSA has few parameters and operators and a simple structure, when the leader falls into the local optimum, it will cause the followers to have the same reaction, resulting in a decrease in the ability to search for food [17]. This is the foundation of the proposed feature selection method in this study. Embedded approach is a method that combines the wrapper and filter approach to exploit the strengths of each method. Based on the embedded approach method, this study proposes a feature selection method that combines SSA and extended repository (ER) to eliminate redundant and irrelevant features by comparing new solutions with solutions already stored in the extended repository to maintain diversity among the salps. At the same time, the mechanisms for updating new positions of individuals in the swarm are improved using the mRMR method and operators of the GA.

In addition to feature selection, feature classification is also an important stage in bearing fault diagnosis. The best feature subset selected by feature selection is used by the classifier to evaluate the classification accuracy. Machine learning applications include many classifiers, such as k -nearest neighbor (KNN), support vector machine (SVM), random forest (RF), domain adaptation (DA), naive Bayes (NB), and decision tree (DT). KNN is an easy-to-understand and widely used technology in machine learning [22], and SVM can make decisions and has good generalization capabilities in a smaller dataset [23]. Therefore, this study uses two classifiers, KNN and SVM, for application.

From the above analysis, a bearing fault diagnosis model is proposed with three main parts. A new feature selection method is proposed based on binary salps combined with an extended repository (BSSA-ER) and new mechanisms to update the positions of individuals in the swarm to find an effective feature subset in the subsequent classification process. The core goal of this study is to develop a model for bearing fault diagnosis. The key points are explained as follows:

1. In the feature extraction stage, three feature extraction methods, MRA, FFT, and EA, are used to effectively extract features from the original signal to form a feature set.
2. A proposal for a feature selection method for BSSA with an extended repository, which is beneficial to maintaining racial diversity.
3. This study validates the proposed model through four datasets. The results show that the SVM classifier is more suitable for this model than the KNN classifier.

2. Feature Extraction Methods

This section introduces three feature extraction methods for extracting time domain and frequency domain features from induction motor current signals, namely EA, MRA, and FFT. The feature extraction process is shown in Figure 1. The original signal was analyzed using methods EA and MRA. Ten statistical feature parameters in the time domain are calculated from the upper envelope and lower envelope. There are 50 statistical features extracted from approximation coefficients and detail coefficients of the MRA. In addition, there are 50 statistical features in the frequency domain calculated from the spectrum of the approximation coefficients and detail coefficients. Table 1 shows the feature statistics used by the characteristic parameters in the time domain and frequency domain, which are maximum value (Max), minimum value (Min), mean value (Mean), mean square error (Mse), and standard deviation (Std), respectively.

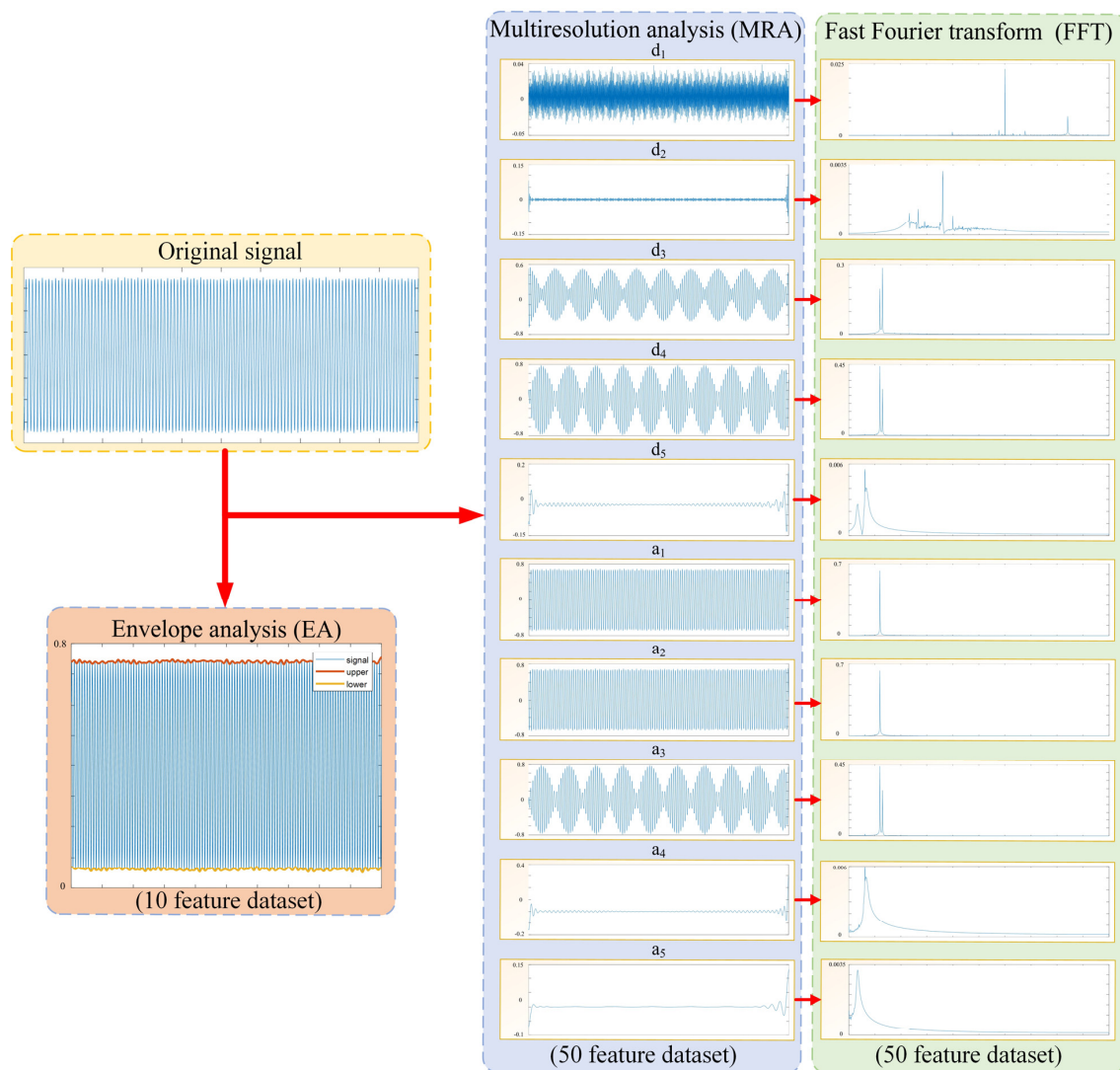


Figure 1. The process of feature extraction using MRA, FFT, and EA.

Table 1. Mathematical expression for each feature in feature selection.

Features	MRA	FFT	EA
	d ₁ d ₂ d ₃ d ₄ d ₅	d ₁ d ₂ d ₃ d ₄ d ₅	
	a ₁ a ₂ a ₃ a ₄ a ₅	a ₁ a ₂ a ₃ a ₄ a ₅	
Max	F1, F2, F3, F4, F5 F6, F7, F8, F9, F10	F11, F12, F13, F14, F15 F16, F17, F18, F19, F20	F101, F102,
Min	F21, F22, F23, F24, F25 F26, F27, F28, F29, F30	F31, F32, F36, F34, F35 F36, F37, F38, F39, F40	F103, F104,
Mean	F41, F42, F43, F44, F45 F46, F47, F48, F49, F50	F51, F52, F53, F54, F55 F56, F57, F58, F59, F60	F105, F106,
Mse	F61, F62, F63, F64, F65 F66, F67, F68, F69, F70	F71, F72, F73, F74, F75 F76, F77, F78, F79, F80	F107, F108,
Std	F81, F82, F83, F84, F85 F86, F87, F88, F89, F90	F91, F92, F93, F94, F95 F96, F97, F98, F99, F100	F109, F110

2.1. Multi-Resolution Analysis

The primary concept of MRA involves identifying the coefficient vector of the discrete wavelet transform (DWT) to analyze the original signal [24]. The original signal can be decomposed by the scaling functions, $\varphi(t)$, and the wavelet functions, $\psi(t)$, into several approximation coefficients, A_j , and detail coefficients, D_j , as expressed by (1), where $\varphi(t)$ and $\psi(t)$ are defined as (2) and (3):

$$x(t) = \sum_k A_{j_0,k} \varphi_{j_0,k}(t) + \sum_j \sum_k D_{j,k} \psi_{j,k}(t) \quad (1)$$

$$\varphi_{j,k}(t) = \sqrt{2^j} \varphi(2^j t - k) \quad (2)$$

$$\psi_{j,k}(t) = \sqrt{2^j} \psi(2^j t - k) \quad (3)$$

Choosing the correct mother wavelet is crucial as it has a significant impact on the usefulness of the extracted signal features [25,26]. This study uses detail coefficients d_1 to d_5 for feature extraction.

2.2. Fast Fourier Transform

In Discrete Fourier Transform (DFT) engineering, FFT is an implementation method [27] in which the calculation process is simplified. It is easier to obtain frequency domain signals by performing bearing time domain FFT transformation. Therefore, frequency domain signals are more suitable for fault diagnosis than original signals. [28] In order to acquire the frequency domain signal, FFT analyzes the time domain signal, $x(t)$, and converts it into the frequency domain, f , of a limited number of frequency lines [29]. The calculation process of FFT is shown in Equation (4). The detail coefficients d_1 to a_5 represent the input of the FFT, as shown in Figure 1.

$$Y(k) = \sum_{n=0}^{N-1} e^{-j2\pi fn/N} x(t) \quad (4)$$

2.3. Envelope Analysis

In envelope analysis, upper and lower envelopes extract the maximum and minimum boundaries of a signal and identify peaks and valleys to create upper and lower envelopes.

These envelopes represent the overall change in signal amplitude. The upper envelope marks the peak amplitude, and the lower envelope represents the troughs.

3. The Feature Selection Method

This section proposes a feature selection method—BSSA-ER. This method is based on two operating mechanisms: (1) updating the salp position according to the feature weight;

and (2) updating the salp position based on genetic calculation. The main goal is to remove redundant features to improve the classification accuracy of the model.

3.1. Binary Salp Swarm Algorithm

Feature selection is performed by converting features into 0 and 1. In other words, each salp position value changes from 1 to 0 to represent that the corresponding feature is not selected, and from 0 to 1 to represent that the corresponding feature is selected. Therefore, the position of each salp is converted into a binary during initialization. The equation is shown in (5).

$$\vec{X}_{id}^{t+1} = \begin{cases} 1, & \text{if } rand_i > 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

SSA divides the group into leaders and followers. The role of the leader salps is to lead the followers to search for food, which is the most important link in the entire algorithm. If there is no leader in the group, the rest are followers. The leader's process of searching for food is represented by (6).

$$x_j^1 = \begin{cases} F_j + c_1((ub_j - lb_j)c_2 + lb_j), & c_3 \geq 0.5 \\ F_j - c_1((ub_j - lb_j)c_2 + lb_j), & c_3 < 0.5 \end{cases} \quad (6)$$

x_j^1 and F_j in the j th dimension represent the positions of the leader and food source, ub_j is the upper limit of the j th dimension, and lb_j is the lower limit of the j th dimension. c_2 and c_3 are randomly selected numbers in the range $[0, 1]$. c_1 is an indispensable parameter in the algorithm, used to balance exploitation and exploration [30]. The equation is shown as (7).

$$x_j^1 = c_1 = 2e^{-(\frac{t}{T})^2} \quad (7)$$

where t represents the current iteration, and T is the maximum number of iterations. The position update equation of the follower salp is shown in (8)

$$x_j^1 = x_j^i = \frac{1}{2}(x_j^i + x_j^{i-1}) \quad (8)$$

3.2. Updating the Salp's Position Based on Feature Weights

This study proposes a salp position update method based on feature weights, using max-relevance and min-redundancy (mRMR) to calculate the correlation. This algorithm considers the correlation of features with class labels and feature redundancy between selected features [31]. Before proceeding, it is necessary to confirm the Hamming distance, S , between the current position of the salp and the group's optimal solution. Hamming distance refers to the proportion of positions where two binary vectors are different. Calculating the Hamming distance between two vectors involves two steps: (1) calculating the two binary vectors using XOR, and (2) calculating the number of 1s in the resulting vector [32]. If half of the Hamming distance is greater than the difference threshold, θ , the update method is started. In this method, the θ threshold is set to one-quarter of the feature length. If the condition $S/2 < \theta$ is reached, it means that the current position of the salp is approaching F , and the different positions of the salp will be adjusted according to their weight considerations. In roulette selection, if a feature has a higher weight, the probability of being selected is higher, and there is a certain chance that a feature with a lower weight will be selected. This is to ensure racial diversity.

3.3. Updating the Slap Position Based on Genetic Operators

Genetic algorithms are computational models of the biological evolution process, which imitates Darwinian biological evolution selection and genetic mechanisms to find the optimal solution [33]. Crossover operation can search new areas in space search, thereby enhancing search capabilities, maintaining the diversity of the group, and avoiding falling

into local optimality. Genetic operators randomly generate a solution race at the beginning, select two solutions as parents, and generate new solutions through crossover mutation. The quality of the parent will affect whether a good solution is produced. To exploit the advantages of GA, this study applies two GA operators, crossover and mutation, in the position updating process. Each location is considered an optimal solution. This will increase the diversity of the population, meaning the algorithm can extend the search space, combined with a refinement process using extended memory to eliminate new locations that overlap with old values. This method is expected to bring greater efficiency in the process of finding the best optimal solution. In addition, this study uses a three-point crossover operation to enable the roulette wheel to select the parent., cutting the solutions at three random points and swapping them with each other to get two new solutions. The mutation operator works by making random changes to the solution, which helps generate new offspring and avoids the population being too similar.

3.4. The Proposed BSSA-ER Method

This study proposes a feature selection method—BSSA-ER. The main purpose of this method is to store all known solutions in the repository. When iterating, the ER stores the solutions, which will be compared with each new solution of the salp. If the solutions in the repository are the same as the new location of the salp, the update mechanism will be triggered. During the process, the new solutions will be confirmed using the extended repository to ensure that the new solutions the approach differs from previously evaluated solutions to avoid excessive computational costs. In addition, combining BSSA and ER will effectively avoid the problem of falling into local optimality and improve the search capability of this method. To avoid falling into an infinite loop, setting the ρ threshold improves the proposed method. Because a dataset has fewer features, the number of solutions will be relatively small, resulting in more time to confirm suitable solutions. In this study, all datasets use $\rho = 100$ to verify the effectiveness of the proposed method.

This study uses a k -nearest neighbor classifier, the number of nearest neighbors, $k = 3$, and 10-fold cross-validation to evaluate each salp. Nearest neighbors, k , is calculated using Euclidean distance. The equation is shown in (9).

$$fitness = \frac{N_{True}}{N_{True} + N_{False}} \times 100\% \quad (9)$$

Finally, the proposed method, BSSA-ER, is completed. The flow chart of BSSA-ER is shown in Figure 2, and the details are as follows:

- Step 1:** Set the parameters (the number of salp N , maximum iteration T).
- Step 2:** Initialize the population of $\vec{X}_i$ in the salp.
- Step 3:** Construct an extended repository, ER.
- Step 4:** Evaluate the fitness value of $\vec{X}_i$ initial salp swarm.
- Step 5:** Obtain the food sources, F , from the salp swarms.
- Step 6:** Update the position of the leader's salp.
- Step 7:** Update the position of the followers' salp.
- Step 8:** Check whether half the Euclidean distance is less than the difference threshold θ . If "Yes", go to **Step 9** for feature weighting. Otherwise, go to **Step 10**.
- Step 9:** Update salp position $\vec{X}_i$ based on feature weights.
- Step 10:** Check if the salp is the same as the solution in the augmentation repository. If "Yes", go to **Step 11**. Otherwise, go to **Step 14**.
- Step 11:** Perform crossover operation.
- Step 12:** Check if the salp is the same as the solution in the augmentation repository. If "Yes", go to **Step 13**. Otherwise, go to **Step 15**.
- Step 13:** Check whether the number of program runs exceeds the ρ threshold. If "Yes", go to **Step 14** for mutation operation. Otherwise, go back to **Step 11**.
- Step 14:** Perform mutation operation.

Step 15: Update the ER and re-evaluate the suitability of the salps in their new position.

Step 16: Check whether t has reached the maximum number of iterations. If “Yes”, end the iteration. If “No”, iteration will continue.

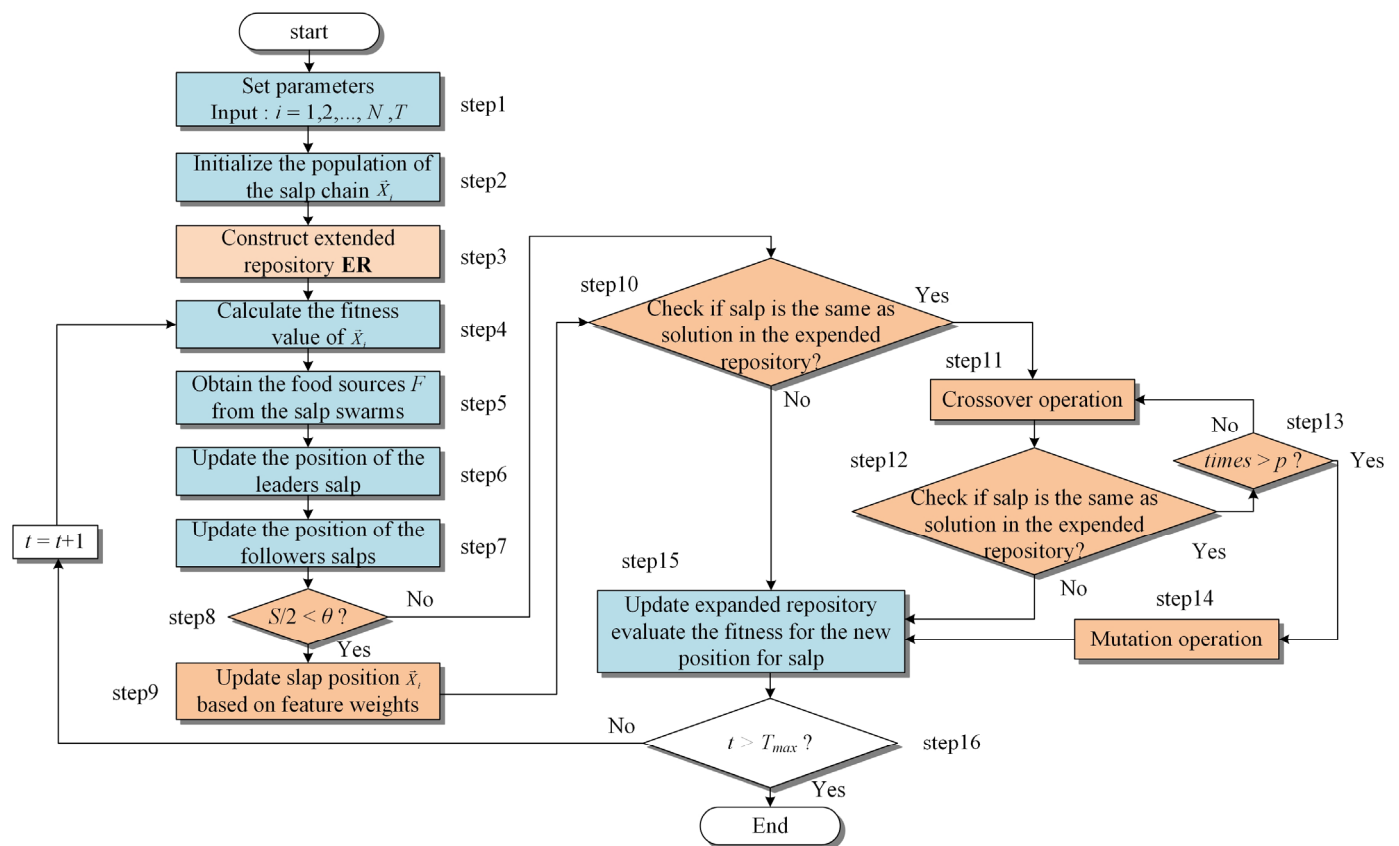


Figure 2. The flow chart of the proposed BSSA-ER.

4. The Bearing Fault Diagnosis Model

This study proposes a motor fault diagnosis model, which can be divided into three stages: obtaining the original signal through feature extraction, then finding the best feature subset through feature selection, and finally using a classifier to evaluate the classification accuracy. The model process is shown in Figure 3. The three stages are explained below:

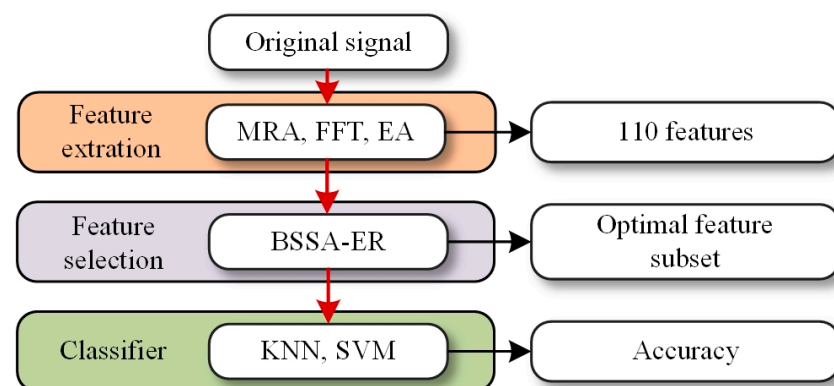


Figure 3. Bearing diagnostic model architecture.

Stage 1: Feature extraction. At this stage, the three methods of MRA, FFT, and EA are combined to extract features from the original signal: MRA extracted 50 features, FFT

extracted 50 features, and EA extracted 10 features. Finally, a total of 110 features were captured in this stage.

Stage 2: Feature selection. At this stage, a fault diagnosis model, BSSA-ER, is proposed. This method can effectively filter out redundant features and select the optimal feature subset, thereby improving classification accuracy.

Stage 3: Classifier. At this stage, two classifiers, KNN and SVM, are used. The KNN is one of the machine learning methods. After calculating the distance using the Euclidean distance, it determines the constant k value for voting to evaluate to which category the k (nearest neighbors) belong. SVM is often used in classification research to find a hyperplane solution. The hyperplane has a maximum margin and is located between two different categories. Finding the hyperplane solution means finding the optimal solution [34].

5. Experimental Verification and Results

To verify the proposed motor fault diagnosis method, this study uses four different datasets and divides them into four cases for explanation, which are explained as follows.

Case study 1: The UCI benchmark dataset is used to verify the search capability of the proposed feature selection method. The proposed method is compared with three different feature selections, namely BSSA, BWOA, and BGWO.

Case study 2: In this case, the motor bearing dataset is selected to verify the search capability of the model. The original current is obtained from the induction motor. The comparison method is the same as case study 1.

Case study 3: The search capability of the model is verified using the benchmark dataset provided by CWRU. In this case, other published methods are used to compare with the proposed fault diagnosis method.

Case study 4: To verify the robustness of the proposed model, this case uses the benchmark dataset provided by the MFPT Association and compares it with other known methods.

Figure 4 briefly describes the process of evaluating the effectiveness of the proposed model and its importance.

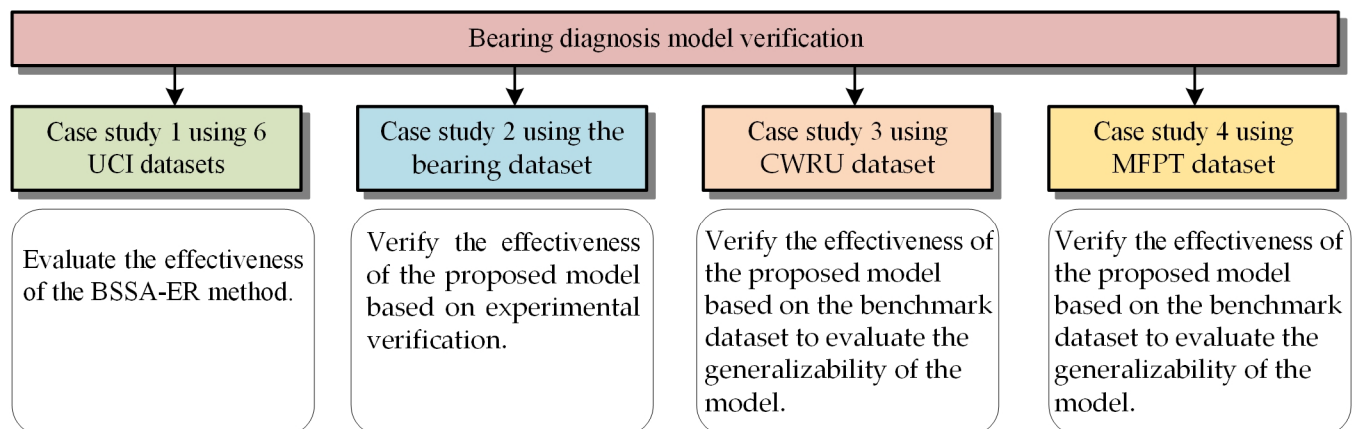


Figure 4. Experimental verification diagram of the bearing diagnosis model.

5.1. Case Study 1: UCI Benchmark Datasets

5.1.1. Description of the Datasets

UCI datasets can be used for training and evaluation of motor fault diagnosis models, so this case study uses six UCI datasets, namely BrestEW, WaveformEW, Sonar, Vote, Vehicle, and HeartEW. The detailed information on the 6 UCI datasets is shown in Table 2.

Table 2. The message of the UCI benchmark datasets.

Datasets	Features	Instances	Classes
BreastEW	30	569	2
WaveformEW	40	5000	2
Sonar	60	208	2
Vote	16	300	2
Vehicle	18	94	5
HeartEW	13	270	2

In the feature selection method using the optimization algorithm, the solution is a sequence of binary numbers, where bit 1 represents selected features and bit 0 stands for unselected features. The length of the binary string is equal to the total number of original features. Therefore, in this study, the optimal solution is the position of the individuals in the swarm. As positions change, new solutions will emerge. The evaluation of solutions is performed by the fitness function of the optimization algorithm. Table 3 details the best optimal feature set of BSSA-ER on the six UCI datasets. The criteria for selecting the best optimal feature subset are based on the best accuracy value (fitness value) and the lowest number of selected features. For the BreastEW dataset, BSSA-ER selected 16 features. This means that the proposed feature selection method can remove 46.6% of the redundant features compared to the original feature sets, WaveformEW, Sonar, Vote, Vehicle, and HeartEW: 32.5%, 41.6%, 37.5%, 38.8%, and 46.2%, respectively.

Table 3. Details of the best optimal feature subset of BSSA-ER on 6 UCI datasets.

Datasets	Best Accuracy	Number of Feature	Feature Index
BreastEW	95.42%	16	F2, F3, F6, F11, F12, F15, F16, F19, F21, F23, F25, F26, F27, F28, F29, F30
WaveformEW	83.66%	27	F1, F2, F3, F4, F5, F6, F7, F8, F9, F10, F11, F12, F13, F14, F15, F16, F17, F19, F21, F29, F30, F31, F36, F37, F38, F39, F40
Sonar	87.98%	35	F1, F3, F4, F7, F9, F10, F11, F12, F13, F15, F16, F22, F23, F24, F26, F30, F31, F32, F33, F34, F37, F38, F39, F43, F44, F46, F48, F49, F50, F51, F54, F55, F58, F59, F60
Vote	95.67%	10	F3, F4, F5, F7, F8, F9, F11, F12, F13, F14
vehicle	73.4%	11	F1, F2, F3, F5, F6, F7, F9, F10, F14, F15, F18
HeartEW	86.67%	7	F3, F7, F9, F10, F11, F12, F13

5.1.2. Comparison with Other Feature Selection Methods

The results of comparing the proposed method with BSSA, BGWO, and BWOA are shown in Table 4. The tables are evaluated based on three criteria: accuracy, number of features, and operating time. Each experiment is executed 30 times. Although the fitness value of the proposed method is 0.59% lower than BWOA in sonar, it is higher than other methods in the datasets of BreastEW, WaveformEW, vote, and vehicle. Among them, the dataset of HeartEW is 2.9% higher than BSSA and is the best.

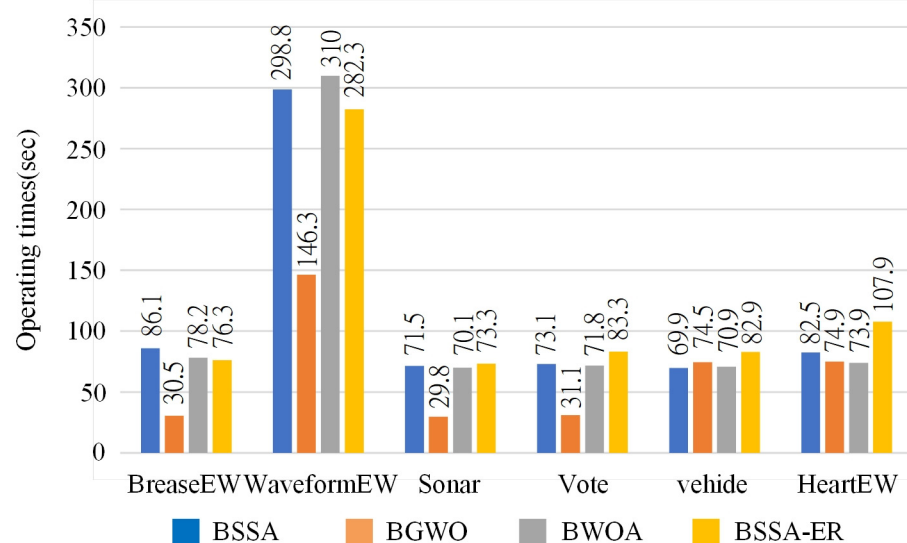
Table 4. Compare with three other methods.

Datasets	BSSA		BGWO		BWOA		The Proposed Method	
	Avg. Fit. (%)	Avg. No. F.	Avg. Fit. (%)	Avg. No. F.	Avg. Fit. (%)	Avg. No. F.	Avg. Fit. (%)	Avg. No. F.
BreastEW	95.04	18.75	95.09	20.35	95.05	12.65	95.23	19.4
WaveformEW	82.85	32.4	83.21	31.2	82.91	33.55	83.35	28.7
Sonar	86.01	40.65	81.92	40.15	87	34.5	86.41	40.56
Vote	95	9.5	95.38	9.45	95.13	7	95.66	10
vehicle	66.76	11.65	68.3	11.4	67.66	8.4	68.9	11.63
HeartEW	82.78	7.3	84.52	7.66	83.15	5.9	85.68	7.33

For each dataset, the results of the better algorithm are bolded. Avg. Fit.: Average Fitness; Avg. No. F.: Average Number of Feature.

The BSSA-ER uses WaveformEW, sonar, and vehicle datasets in the feature number part, which are 3.7, 0.09, and 0.02 less than BSSA, respectively. Among them, the WaveformEW dataset is even smaller than the other three methods. However, the BreastEW, Vote, and HeartEW datasets are 0.65, 0.5, and 0.03 more than BSSA, respectively. Although the proposed method failed to reduce the average number of selected features in terms of the average number of features, it has the best average accuracy.

Finally, although the proposed method failed to achieve the best results in terms of running time, it was 9.8 s, 16.5 s, 1.9 s, and 27.7 s lower than BSSA and BWOA on the BreastEW and WaveformEW datasets, respectively. Figure 5 shows a comparison of the average operating times of the methods using UCI data. In conclusion, the proposed method has better search capabilities than the other three methods.

**Figure 5.** Compare average operating times of methods using UCI data.

5.2. Case Study 2: Motor Dataset of Current Signal

5.2.1. Description of the Dataset

The original signal obtained from the induction motor was used in this case study. Figure 6 shows the laboratory equipment, which included an induction motor (2 HP, 4 poles, 380 V, 60 HZ), a data miner (NI PXI-1037), a servo motor (FUKUTA), an oscilloscope, and a personal computer. In addition, the hardware specification of the computer used in this work was a personal computer with Windows 10 operating system, Intel(R) Core i5-6500 CPU @ 3.20 GHz, and 16.0 GB RAM. The software used in the resulting simulation was MATLAB 2017-a version. The original signal measured from the motor is affected by the environment, so it is usually full of noise. To conform to the actual situation encountered,

this case adds different levels of Gaussian white noise to the original signal. Signal-to-noise ratio (SNR) is a standard for noise levels, with ∞ dB representing no noise and 0 dB representing maximum noise. Figure 7 shows a bearing inner ring square hole failure. There are three different fault types of square holes in the bearing inner ring. The inner ring square hole comprises three fault sizes, namely $0.53 \text{ mm} \times 0.53 \text{ mm}$, $0.53 \text{ mm} \times 1.24 \text{ mm}$, and $0.53 \text{ mm} \times 1.96 \text{ mm}$.

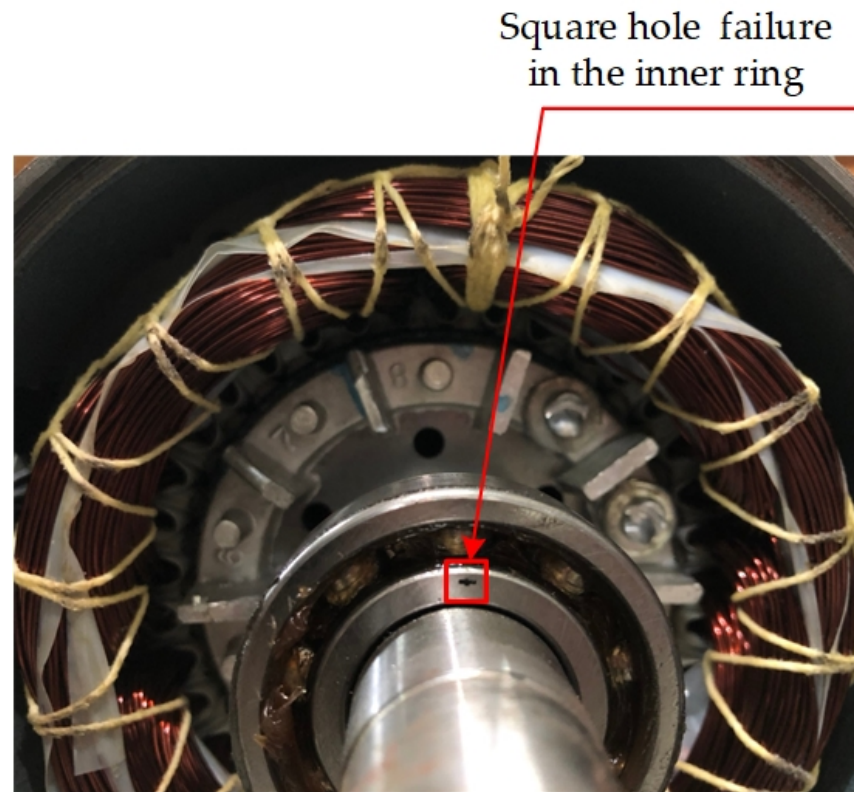


Figure 6. Square hole failure in the inner ring of the bearing.

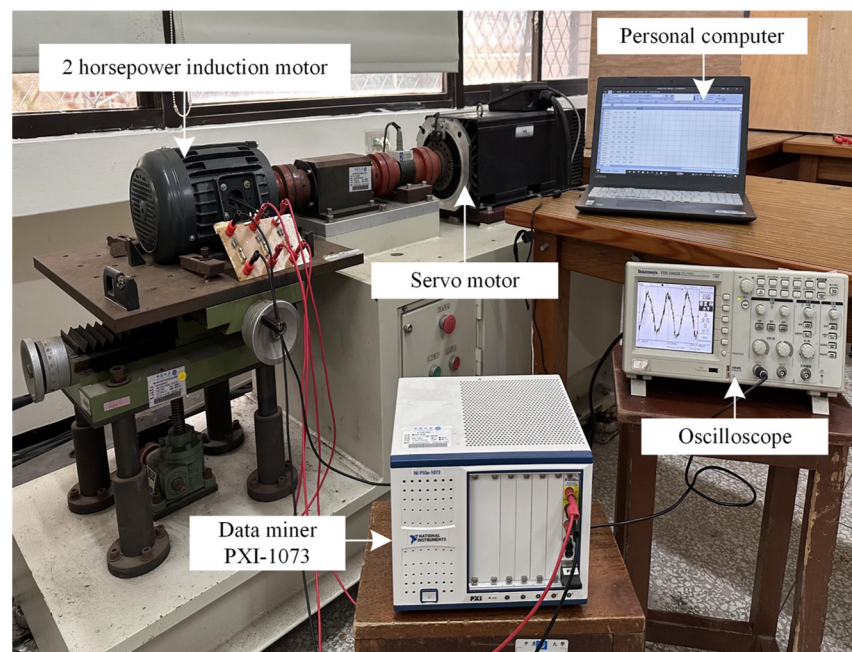


Figure 7. The laboratory equipment.

5.2.2. Comparison with Other Feature Selection Methods

The best feature subsets after feature selection are stored in the dataset, and then to KNN and SVM, respectively. The purpose of this case is to evaluate the search capabilities and stability of the proposed method. The average accuracy of the KNN classifier is shown in Table 5. The proposed method has the best accuracy of 99.28% under both ∞ dB and 30 dB noise. Under 20 dB noise, the accuracy of the proposed method is higher than that of BSSA (0.03%) and BWOA (0.04%), respectively, and only lower than BGWO (0.17%). Under 10 dB noise, the accuracy of the proposed method is higher than BSSA (0.17%) and BWOA (0.16%), respectively, and only lower than BGWO (0.19%). Under 0 dB noise, the accuracy of the proposed method is higher than BSSA (0.31%) and BWOA (0.01%), respectively, and only lower than BGWO (0.1%). The average accuracy of the SVM classifier is shown in Table 6. Under 30 dB, 20 dB, and 0 dB noise, the proposed method has the best results, which were 92.05%, 79.4%, and 55.6%, respectively. Under ∞ dB noise, the accuracy of the proposed method is only 0.05% lower than BWOA. Under 10 dB noise, the accuracy of the proposed method is only 0.3% lower than BGWO. This shows the robustness of the proposed method, except for noise of 10 dB. Tables 5 and 6 compare the results. The results of the employed SVM classifier under different noise conditions are better than those of the employed KNN classifier. Therefore, this case is suitable for using the SVM classifier. Figure 8 shows the average operating times of the four methods under different noise levels. Among all the methods, BGWO has the longest operating time on every dataset. The average operation time of BSSA-ER is less than 8.4 s, 9.6 s, 5.2 s, 10.9 s, and 2.6 s than that of BSSA, which means that this method takes less time to calculate and reduces the calculation cost compared to BSSA. In this case study, the feasibility of the proposed fault diagnosis model is verified; in addition, according to the experimental results, the SVM classifier is more applicable for the study of this case than the KNN classifier.



Figure 8. Average operating times by comparison methods.

Table 5. The average accuracy by using the KNN classifier.

Methods	Average Accuracy (%)				
	∞ dB	30 dB	20 dB	10 dB	0 dB
BGWO	97.5	89.94	78.27	64.95	53.45
BWOA	99.68	89.59	77.98	64.6	53.34
BSSA	96.98	90.03	77.97	64.59	53.04
Proposed	99.28	90.44	78.1	64.76	53.35

The results of the better algorithm are bolded.

Table 6. The average accuracy by using the SVM classifier.

Methods	Average Accuracy (%)				
	∞ dB	30 dB	20 dB	10 dB	0 dB
BGWO	99.85	92.9	79	65.4	54.95
BWOA	99.95	92.05	79.35	65	53.85
BSSA	99.8	92.8	79.2	63.75	55.55
Proposed	99.9	92.95	79.4	65.1	55.6

The results of the better algorithm are bolded.

5.3. Case Study 3: CWRU Benchmark Dataset

5.3.1. Description of the Dataset

This study uses the CWRU benchmark dataset to verify the proposed fault diagnosis model. CWRU is a benchmark dataset provided by Case Western Reserve University. Its experimental equipment includes a 2 HP induction motor, a load motor, and a torque sensor. Test motor faults are divided into three types: inner race, outer race, and outer race. Each bearing failure is EDM-machined into three different sizes: 0.007-inch, 0.014-inch, and 0.021-inch in diameter. This study divides this dataset into ten categories, one category under normal circumstances, and three categories each for the inner circle, outer circle, and sphere. The number of samples in each class is 100, with a length of 1200 sample points for each sample. The detailed information and classification are shown in Table 7.

Table 7. CWRU benchmark dataset classification method.

Fault Location	Loads (hp)	Defect Diameters (Inches)	Samples	Classes
Normal	0		100	1
Inner race	0	0.007	100	2
		0.014	100	3
		0.021	100	4
Ball	0	0.007	100	5
		0.014	100	6
		0.021	100	7
Outer race	0	0.007	100	8
		0.014	100	9
		0.021	100	10

5.3.2. Comparison with Other Feature Selection Methods

Li et al. proposed a random window sampling (RWS) method to extend the dataset to avoid data and overfitting [35]. Li et al. applied LMD to decompose the bearing vibration signal into multiple product functions (PF) and then extracted the multi-scale permutation entropy (MPE) of the main PF. To complete automatic fault classification, an improved binary tree based on a support vector machine was proposed (ISVM-BT) [36]. An et al. proposed a supervised feature extraction method based on sparse expression and mapping sample features to feature domains through effective transformation [37]. In this case, the proposed model is compared with models published in the last five years. The proposed models use KNN and SVM classifiers, respectively. The results are shown in Table 8. The proposed model achieved 99.5% accuracy using KNN and 99.4% accuracy using SVM. Although in the comparison of accuracy, the results obtained by the other three methods are higher than the proposed model using KNN and SVM, they are better than the accuracy of FMCNN using SVM. The results show that the proposed model has good classification ability and is suitable for motor fault diagnosis.

Table 8. The proposed model compared with the known studies using the CWRU dataset.

Models	Classes	Accuracy (%)	Public Years
RWS [28]	10	99.9	2023
ISVM-BT [29]	10	99.9	2020
FMCNN [30]	10	98.8	2019
BSSA-ER-KNN	10	99.5	
BSSA-ER-SVM	10	99.6	

5.4. Case Study 4: MFPT Benchmark Dataset

5.4.1. Description of the Dataset

The MFPT used in this case is a benchmark dataset commonly used for fault diagnosis, including three different fault types: baseline condition, outer ring fault condition, and inner ring fault condition. The input shaft speeds were all 25 Hz. The baseline load was 270 pounds, and vibration data were captured at a sampling rate of 97,656 Hz for 6 s. Vibration data were captured for outer ring failure and inner ring failure at a sampling rate of 48,828 Hz for 3 s, with loads of 0, 50, 100, 150, 200, 250, and 300 pounds, respectively. The data were finally classified into three classes. The details of the MFPT benchmark dataset are shown in Table 9.

Table 9. MFPT benchmark dataset classification method.

Fault Location	Load (Pound)	Samples	Classes
Baseline	270	200	1
Inner race	0/50/100/150/200/250/300	350	2
Outer race	0/50/100/150/200/250/300	350	3

5.4.2. Comparison with Other Feature Selection Methods

In this case, the proposed motor fault diagnosis model is compared with models published in the last five years. Li et al. proposed a network architecture search method (RL-NAS) with fault diagnosis ability, search space, search efficiency, and multi-working condition performance [38]. Yuan et al. proposed a deep neural network model based on a combination of CNN and SVM for the fault diagnosis of rolling bearings [1]. Sun et al. used Gaussian modulated linear group delay (GLGD) to propose a second-order time redistribution multi-synchronous squeezing transform (STMSST) to obtain time–frequency images with high resolution [39]. Table 10 shows the comparison results between the proposed model and the other three fault diagnosis methods. The accuracy of this model using the KNN classifier is 98.45%, and the accuracy using the SVM classifier is 98.98%. The proposed method uses the SVM classifier more effectively than the other three methods. From this case, the SVM classifier is more applicable for use in this study.

Table 10. The proposed model is compared with known studies using the MFPT dataset.

Models	Classes	Accuracy (%)	Public Years
RL-NAS [31]	3	96.86	2023
CNN-SVM [32]	3	98.89	2020
STMSST [33]	3	98.67	2020
BSSA-ER-KNN	3	98.45	
BSSA-ER-SVM	3	98.98	

In summary, the proposed model combined with the SVM classifier is a simple, highly effective, and competitive diagnostic model compared to the state-of-the-art models. However, some existing problems of the proposed model need to be extended and studied further, such as calculation time and assessment of the overfitting problem when using

SVM. In addition, the process of building the input dataset for this study also depends on the experience of the expert.

6. Conclusions

This study proposes an efficient model that uses a combination of three feature extraction methods: MRA, FFT, and EA to extract 110 features from the original signal. In the feature selection stage, a method combining BSSA-ER is proposed to eliminate redundant and irrelevant features to achieve more effective feature selection. Finally, KNN and SVM classifiers were used to classify the selected features for feature selection, and then the four datasets were divided into four cases to evaluate the search capability and stability of the proposed model. In case study 1, the UCI benchmark dataset was used to evaluate the effectiveness of the proposed model. Compared with the other three methods, the proposed method had better search capability. The results show that it is highly competitive compared to traditional techniques and can be applied in many fields. Case study 2 used the original signal obtained by the induction motor to evaluate the proposed model. To conform to the actual situation encountered, Gaussian white noise was added. The proposed method used the SVM classifier to obtain the best accuracy under different white noise levels and had excellent running times in both the BSSA and BGWO methods. In case study 3, the CWRU benchmark dataset was employed, comparing the proposed model with four other models. The proposed model obtained 99.4% accuracy by applying the SVM classifier. In case study 4, using the MFPT quantity set to compare with the other three models, the proposed method obtained an accuracy of 98.98% using the SVM classifier. To summarize, the proposed model achieved better results using the SVM classifier than using the KNN classifier. In addition, the competitiveness of the proposed model was compared with previously published models using the same public dataset. Therefore, the proposed model can be applied in motor fault diagnosis, but there is still room for improvement in running time.

Author Contributions: Methodology, C.-Y.L., T.-A.L., Y.-C.C. and S.-C.H.; visualization, C.-Y.L. and Y.-C.C.; software, C.-Y.L.; data curation, Y.-C.C.; writing—original draft preparation, Y.-C.C.; resources, C.-Y.L.; investigation, T.-A.L.; writing—review and editing, C.-Y.L., T.-A.L. and S.-C.H. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Science and Technology Council under grant No. NSTC 112-2221-E-033-008-MY3.

Data Availability Statement: The data will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Yuan, L.; Lian, D.; Kang, X.; Chen, Y.; Zhai, K. Rolling bearing fault diagnosis based on convolutional neural network and support vector machine. *IEEE Access* **2020**, *8*, 137395–137406. [\[CrossRef\]](#)
2. Li, C.; de Oliveira, J.L.V.; Lozada, M.C.; Cabrera, D.; Sanchez, V.; Zurita, G. A systematic review of fuzzy formalisms for bearing fault diagnosis. *IEEE Trans. Fuzzy Syst.* **2018**, *27*, 1362–1382. [\[CrossRef\]](#)
3. Wang, H.; Xu, J.; Yan, R.; Gao, R.X. A New Intelligent Bearing Fault Diagnosis Method Using SDP Representation and SE-CNN. *IEEE Trans. Instrum. Meas.* **2019**, *69*, 2377–2389. [\[CrossRef\]](#)
4. Lee, C.Y.; Le, T.A. Intelligence Bearing Fault Diagnosis Model Using Multiple Feature Extraction and Binary Particle Swarm Optimization with Extended Memory. *IEEE Access* **2020**, *8*, 198343–198356. [\[CrossRef\]](#)
5. Huang, Z.; Wang, Z.; Zhang, H. Multilevel feature moving average ratio method for fault diagnosis of the microgrid inverter switch. *IEEE/CAA J. Autom. Sin.* **2017**, *4*, 177–185. [\[CrossRef\]](#)
6. Brigham, E.O.; Morrow, R.E. The fast Fourier transform. *IEEE Spectr.* **1967**, *4*, 63–70. [\[CrossRef\]](#)
7. Randall, R.B.; Antoni, J.; Chobsaard, S. The Relationship between Spectral Correlation and Envelope Analysis in the Diagnostics of Bearing Faults and Other Cyclostationary Machine Signals. *Mech. Syst. Signal Process.* **2001**, *15*, 945–962. [\[CrossRef\]](#)
8. Ghelardoni, L.; Ghio, A.; Anguita, D. Energy Load Forecasting Using Empirical Mode Decomposition and Support Vector Regression. *IEEE Trans. Smart Grid* **2013**, *4*, 549–556. [\[CrossRef\]](#)

9. Huang, E.N.; Shen, S.S.P. *Hilbert-Huang Transform and Its Applications*; World Scientific Pub. Co Pte Ltd.: Chennai, India, 2005; Volume 16.
10. Peng, Z.; Chu, F. Application of the wavelet transform in machine condition monitoring and fault diagnostics: A review with bibliography. *Mech. Syst. Signal Process.* **2004**, *18*, 199–221. [\[CrossRef\]](#)
11. Devadasu, G.; Sushama, M. A novel multiple fault identification with fast Fourier transform analysis. In Proceedings of the 2016 International Conference on Emerging Trends in Engineering, Technology and Science (ICETETS), Pudukkottai, India, 24–26 February 2016.
12. Yang, Y. A signal theoretic approach for envelope analysis of real-valued signals. *IEEE Access* **2017**, *5*, 5623–5630. [\[CrossRef\]](#)
13. Lee, C.-Y.; Le, T.-A.; Hung, C.-L. A feature selection approach based on memory space computation genetic algorithm applied in bearing fault diagnosis model. *IEEE Access* **2023**, *11*, 51282–51295. [\[CrossRef\]](#)
14. Hall, M.A. *Correlation-Based Feature Selection for Machine Learning*; The University of Waikato: Hamilton, New Zealand, 1999.
15. Abualigah, L.; Shehab, M.; Alshinwan, M.; Alabool, H. Salp swarm algorithm: A comprehensive survey. *Neural Comput. Appl.* **2019**, *32*, 11195–11215. [\[CrossRef\]](#)
16. Wang, D.; Tan, D.; Liu, L. Particle swarm optimization algorithm: An overview. *Soft Comput.* **2018**, *22*, 387–408. [\[CrossRef\]](#)
17. Sulaiman, M.H.; Mustaffa, Z.; Mohamed, M.R.; Aliman, O. Using the Gray Wolf Optimizer for Solving Optimal Reactive Power Dispatch Problem. *Appl. Soft Comput. J.* **2015**, *32*, 286–292. [\[CrossRef\]](#)
18. Mirjalili, S.; Lewis, A. The whale optimization algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [\[CrossRef\]](#)
19. Kumar, M.; Husian, M.; Upreti, N.; Gupta, D. Genetic Algorithm: Review and Application. *Int. J. Inf. Technol. Knowl. Manag.* **2010**, *2*, 451–454. [\[CrossRef\]](#)
20. Mirjalili, S.; Gandomi, A.H.; Mirjalili, S.Z.; Saremi, S.; Faris, H.; Mirjalili, S.M. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Adv. Eng. Softw.* **2017**, *114*, 163–191. [\[CrossRef\]](#)
21. Zhang, J.; Wang, J. Improved Salp Swarm Algorithm Based on Levy Flight and Sine Cosine Operator. *IEEE Access* **2020**, *8*, 99740–99771. [\[CrossRef\]](#)
22. Xing, W.; Bei, Y. Medical Health Big Data Classification Based on KNN Classification Algorithm. *IEEE Access* **2020**, *8*, 28808–28819. [\[CrossRef\]](#)
23. Lee, C.; Le, T.; Lin, Y. A Feature Selection Approach Hybrid Grey Wolf and Heap-Based Optimizer Applied in Bearing Fault Diagnosis. *IEEE Access* **2022**, *10*, 56691–56705. [\[CrossRef\]](#)
24. Félix, C.S.; Walter, F.G.; Raúl, G.C.; Stephane, C.D. Multiresolution analysis based on Mallat pyramidal algorithm applied to GPR data. In Proceedings of the 15th International Conference on Ground Penetrating Radar, Brussels, Belgium, 30 June–4 July 2014; pp. 647–650.
25. Vakharia, V.; Gupta, V.; Kankar, P. A multiscale permutation entropy-based approach to select wavelet for fault diagnosis of ball bearings. *J. Vib. Control* **2014**, *21*, 3123–3131. [\[CrossRef\]](#)
26. He, H.; Tan, Y.; Wang, Y. Optimal Base Wavelet Selection for ECG Noise Reduction Using a Comprehensive Entropy Criterion. *Entropy* **2015**, *17*, 6093–6109. [\[CrossRef\]](#)
27. Zhang, M.; Yin, J.; Chen, W. Rolling Bearing Fault Diagnosis Based on Time-Frequency Feature Extraction and IBA-SVM. *IEEE Access* **2022**, *10*, 85641–85654. [\[CrossRef\]](#)
28. Zhong, Z.; Liu, H.; Mao, W.; Xie, X.; Cui, Y. Rolling Bearing Fault Diagnosis across Operating Conditions Based on Unsupervised Domain Adaptation. *Lubricants* **2023**, *11*, 383. [\[CrossRef\]](#)
29. Strömbergsson, D.; Marklund, P.; Berglund, K.; Larsson, P.E. Bearing monitoring in the wind turbine drivetrain: A comparative study of the FFT and wavelet transforms. *Wind Energy* **2020**, *23*, 1381–1393. [\[CrossRef\]](#)
30. Chen, R.; Dong, C.; Ye, Y.; Chen, Z.; Liu, Y. QSSA: Quantum evolutionary salp swarm algorithm for mechanical design. *IEEE Access* **2019**, *7*, 145582–145595. [\[CrossRef\]](#)
31. Wang, G.; Lauri, F.; Hassani, A.H.E. Feature Selection by mRMR Method for Heart Disease Diagnosis. *IEEE Access* **2022**, *10*, 100786–100796. [\[CrossRef\]](#)
32. Banka, H.; Dara, S. A Hamming distance based binary particle swarm optimization (HDBPSO) algorithm for high dimensional feature selection, classification and validation. *Pattern Recognit. Lett.* **2015**, *52*, 94–100. [\[CrossRef\]](#)
33. Li, J.; Lei, L. A Hybrid Genetic Algorithm Based on Information Entropy and Game Theory. *IEEE Access* **2020**, *8*, 36602–36611.
34. Altayef, E.; Anayi, F.; Packianather, M.; Benmahamed, Y.; Kherif, O. Detection and Classification of Lamination Faults in a 15 kVA Three-Phase Transformer Core Using SVM, KNN and DT Algorithms. *IEEE Access* **2022**, *10*, 50925–50932. [\[CrossRef\]](#)
35. Meng, Z.; Luo, C.; Li, J.; Cao, L.; Fan, F. Research on fault diagnosis of rolling bearing based on lightweight model with multiscale features. *IEEE Sens. J.* **2023**, *23*, 13236–13247. [\[CrossRef\]](#)
36. Yan, X.; Liu, Y.; Jia, M. Multiscale cascading deep belief network for fault identification of rotating machinery under various working conditions. *Knowl.-Based Syst.* **2020**, *193*, 105484. [\[CrossRef\]](#)
37. An, F.P. Rolling bearing fault diagnosis algorithm based on FMCNN-Sparse representation. *IEEE Access* **2019**, *7*, 102249–102263. [\[CrossRef\]](#)

38. Li, J.; Cao, X.; Chen, R.; Zhang, X.; Huang, X.; Qu, Y. Graph neural network architecture search for rotating machinery fault diagnosis based on reinforcement learning. *Mech. Syst. Signal Process.* **2023**, *202*, 110701. [[CrossRef](#)]
39. Sun, G.; Gao, Y.; Xu, Y.; Feng, W. Data-Driven Fault Diagnosis Method Based on Second-Order Time-Reassigned Multisynchrosqueezing Transform and Evenly Mini-Batch Training. *IEEE Access* **2020**, *8*, 120859–120869. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.