

Article

Induction Motor Multiclass Fault Diagnosis Based on Mean Impact Value and PSO-BPNN

Chun-Yao Lee * and Hong-Yi Ou

Department of Electrical Engineering, Chung Yuan Christian University, No. 200, Zhongbei Road, Zhongli District, Taoyuan City 320, Taiwan; g10878005@cycu.edu.tw

* Correspondence: cyl@cycu.edu.tw; Tel.: +886-3-2654827

Abstract: This paper presents a feature selection model based on mean impact value (MIV) to solve induction motor (IM) fault diagnosis on the current signal. In this paper, particle swarm optimization (PSO) is combined with back propagation neural network (BPNN) to classify the current signal of IM. First, the purpose of this study is to establish IM fault diagnosis system. Additionally, this study proposes a feature selection process that is composed of MIV, whose objective is to reduce the number of classifier input features. Secondly, the features are extracted as a feature database after analyzing the current signal of IM, and the fault diagnosis is established through the model of PSO-BPNN. Finally, redundant features are deleted through this feature selection process and a classifier is built. The result shows that the feature selection model based on MIV can filter the features effectively at a signal to noise ratio of 30 dB and 20 dB for the IM fault detection problem. In addition, the computing time of BPNN is also reduced which is helpful for online detection.

Keywords: induction motor; fault diagnosis; back propagation neural network; S-transform



Citation: Lee, C.-Y.; Ou, H.-Y. Induction Motor Multiclass Fault Diagnosis Based on Mean Impact Value and PSO-BPNN. *Symmetry* **2021**, *13*, 104. <https://doi.org/10.3390/sym13010104>

Received: 15 December 2020

Accepted: 6 January 2021

Published: 8 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

In the industry age, machinery is inseparable from our lives, and the core of machinery is divided into motors and engines. With the increasing awareness of environmental protection and the imminent exhaustion of fossil fuels, more and more applications of motors also indicate that industries have increased their dependence on motors. Thus, the losses caused by motor failures and shutdowns will also increase. Therefore, fault detection and diagnosis are important to avoid unexpected shutdowns and degraded efficiency [1]. If the fault of a motor can be identified before shutdown, more time can be arranged for maintenance or replacement, and the loss or danger caused by sudden shutdown can be reduced. Therefore, this paper proposes a feature selection model to delete redundant features of classifier. Moreover, the computing time of classifier can be reduced. Hence, this classifier can be used to detect the fault of the IM faster and more accurately [2].

In another research [3], due to the robustness, low cost, and versatility of motors, they are the electric equipment used in the most various projects in the factory. The motor accounts for 85% of the global energy consumption. The motor is not expected to be subjected to excessive pressure and use. Therefore, the fault diagnosis of the motor has received great attention. According to [3], the most common fault condition is bearing damage, accounting for 41%; followed by stator failure, accounting for 37%; and finally, rotor failure, accounting for 10%. When the motor fails, the running cost will increase. Therefore, fault identification has always been one of the most-discussed topics in industrial applications. The most common technologies associated with motors are temperature measurement, vibration, electrical signal, and sound measurement [4,5]. The current signal used in this paper as a basis for identification is a type of electrical signal. When measuring electrical signals, no additional sensors are needed, and the environmental influence during measurement is also smaller than temperature, vibration, and sound signal. In this study, fault identification is performed for bearing damage, stator short-circuits, and rotor

drilling. This study uses IM current signals for analysis. Because additional sensors for measurement are not required for current signals, and are less affected by environmental noise, they can be used for online real-time monitoring. On-line health monitoring systems play an important role in avoiding unexpected fault and obtaining a higher accuracy for maintenance. Stator current spectra has been well documented on different techniques [6]. In addition, there are noise signals due to working environment of motor is complicated. These noises are mainly Gaussian noise [7].

It has been a challenge for researchers to detect faults of motors [8]. The current signal is regarded as an important detection basis for fault identification. In [8], the current signal was successfully used to detect the broken rotor bar and the bearing failure. The current signal is also used to identify the motors in four conditions in this study. The accuracy of the classifier can be improved through feature engineering [9]. Feature construction, feature extraction, and feature selection are common feature engineering. Feature selection can be divided into two categories, filter and wrapper [10]. The filter involves the correlation of features to select the feature. The time cost of calculation is higher. The wrapper is based on evaluation function to select feature. The time cost of calculation is lower. These tasks can be divided into three directions: data acquisition, feature extraction, and classification. It is the fault detection and diagnosis (FDD) [11].

Time-frequency analysis has greatly developed in the past few decades, which is mostly applied in earth science. Short time Fourier transform (STFT), wavelet transform (WT), and S-transform (ST) are common T-F transformers. Those are the common methods for spectral decomposition [12]. The development of S-transform is based on STFT and WT. S-transform uses a frequency-dependent variable Gaussian function as the window function to overcome the fixed frequency analysis problem of STFT. S-transform also has a phase factor, which is a characteristic not found in wavelet transform [13]. ST has the advantages of STFT and WT, so it is a good choice for time-frequency analysis. Although ST is mainly used in geophysics, it is also used in other fields such as gearbox fault diagnosis, signal recognition and recovery [14–18]. Therefore, in this paper, ST is used to analyze the current signal of the IM in order to find the identification features.

Principal component analysis (PCA) is considered as a method of feature extraction in dimensional reduction for machine learning. Many applications use PCA to extract feature. PCA is characterized by its self-learning ability [19]. However, PCA is sensitive to outliers and missing data. It requires a lot of calculation for calculating the covariance matrix [20]. The mean impact value (MIV) can show the importance of features, and the operation time cost is low. In terms of feature processing, MIV is a more suitable choice considering time cost.

MIV can effectively show the relationship between the input features and output of back propagation neural network (BPNN) [21]. The feature is added to the input features of the neural network in order from the small MIV to the large MIV in [21]. Choosing the best combination of features has good results in predicting blood pressure. PNN-based feature selection (PFS) is a systematic way to select features [22], and the research results in [22] also show the reliability of this process. This study combines the above characteristics to analyze the IM current signal. It is evident that MIV can effectively show the importance of features. In [22], we noticed that the recognition of BPNN is the best. The concept of PFS feature selection is used to build a feature selection model base on MIV and take BPNN as the classifier in this study. This study builds a new feature selection model to simplify the classifier and it can be effectively applied to the fault identification of IM. This feature selection model is backward sequential selection (BSS). In [23], forward sequential selection (FSS) and BSS can improve the detection rate.

This paper proposes a feature selection model to filter out irrelevant features. BPNN is used to detect the fault of IM. First, we get the current signal of IM, and analyze the current signal by S-transform. Second, the feature is extracted in order to detect the fault of IM. Finally, we use the feature selection model which is proposed in this paper to delete irrelevant feature, then the input features and compute time of BPNN can be reduced.

2. Signal Analysis Method and Classifier

2.1. S-Transform (ST)

S-transform is a time-frequency analysis method published in 1994 [24]. The main concept is the extension of continuous wavelet transform, and the operation method is based on Fourier transform. The S-transform analysis method is based on Gaussian window function. This window function will have different window widths due to different frequency. The window width at high frequency becomes smaller and the resolution in the time domain is higher; at a low frequency, the width of the window becomes larger and the resolution in the frequency domain is higher, so it can show the local characteristics of the signal. S-transform compares with the wavelet transform and the Fourier transform, it has a better resolution for the non-steady-state signal and the applicable fields of S-transform are extensive, including high-altitude wind direction monitoring, the detection of gravitational waves, the detection of power transmission network interference, etc [25]. The S-transform formula is shown in Equation (1):

$$S_x(t, f) = \int_{-\infty}^{\infty} x(\tau) |f| e^{-\pi(t-\tau)^2 f^2} e^{-j2\pi f \tau} d\tau \quad (1)$$

The window function is a Gaussian window function as shown in Formula (2):

$$w(t, f) = |f| e^{-\pi t^2 f^2} \quad (2)$$

2.2. Neural Network

In the machine learning, a neural network is a mathematical model created by imitating the biological nervous system. By mimicking the way that the biological nervous system transmits signals, different activation functions are used to transmit the signal to the next neuron; the concept of the computing direction is as following: input the feature vector to the first layer. After calculating the activation function of the first layer, output to the second layer as the input vector of the second layer, and after calculating the activation function of the second layer, output to the third layer as the input vector, which is passed to the output layer and outputs the result.

The calculation method of each neuron is to give the corresponding weights and biases to all input features. After adding the bias and the product of the input features and the weights, the output is performed through the activation function. The simplest activation function is to set a threshold. If the operation result is larger than the threshold, it outputs 1; if the operation result is less than the threshold, it outputs 0. This is used to solve the dichotomy problem. The activation function used in this study is the sigmoid function, which refines the output only 0 and 1 to any real number between 0 and 1.

The core of machine learning is to find the mathematical relationship between the input features and the output through these neurons and modify the weight and the bias from multiple data in the database to optimize the output error. The methods of machine learning are divided into two categories: forward propagation and backward propagation. This paper uses a backward propagation neural network, which is called a BPNN. The BPNN used in this paper has three layers. The activation function between the input layer and the hidden layer is sigmoid function, and the activation function between the hidden layer and output layer is linear function (purelin). We use least squares method to modify the weight. This step is called neural network training. The last layer is the output layer, and there are four neurons in the output layer. These four neurons represent the four types of motors. After the BPNN calculation, the four neurons are used for classification, and the input data will be classified into the category with the largest value among the four neurons.

3. Methodology

Many feature selection methods have been proposed in previous studies. These methods can be divided into two types: filter methods and wrapper methods [26]. The

wrapper method evaluates the importance of the feature set through the results of the classifier. Generally, the effect is better than the filtering method [27]. The feature selection based on MIV used in this study is a wrapper method.

In this study, BPNN is used as the classifier. There are 50 features extracted from the motor current signal after S-transform as identification features. However, some of the 50 features have no positive effect on the fault identification, so these 50 features need to be selected. This section will introduce the feature selection process and method used in this study, which can select several key features from the original 50 features to achieve the effect of simplifying BPNN and computing time.

3.1. Mean Impact Value

In the study [21], the mean impact value (MIV) can effectively show the importance of the feature to the classifier; if MIV of the feature is large, it means that the feature is important for the classification model, and that the feature is more important for classification. This study analyzes the relationship between each feature and the classification result by calculating the MIV of 50 feature values, and uses this analysis result as the basis for the process of feature selection. The process of MIV is shown in Figure 1, and the calculation process steps are as follows:

Step (1) Select all features as feature set $F = \{F_1, F_2, \dots, F_j\}$

Step (2) Train the model of PSO-BPNN.

Step (3) Assume adjustment rate $\pm R$ and adjust F_i , get $F_{i1} = \{F_1, F_2, \dots, F_i(1+R), \dots, F_j\}$ and $F_{i2} = \{F_1, F_2, \dots, F_i(1-R), \dots, F_j\}$.

Step (4) Respectively input F_{i1} and F_{i2} to BPNN.

Step (5) Get the output Y_{i1} and Y_{i2} .

Step (6) Calculate impact value of F_i , $IV_i = Y_{i1} - Y_{i2}$.

Step (7) Calculate mean impact value, $MIV_i = \text{mean}(IV_i)$.

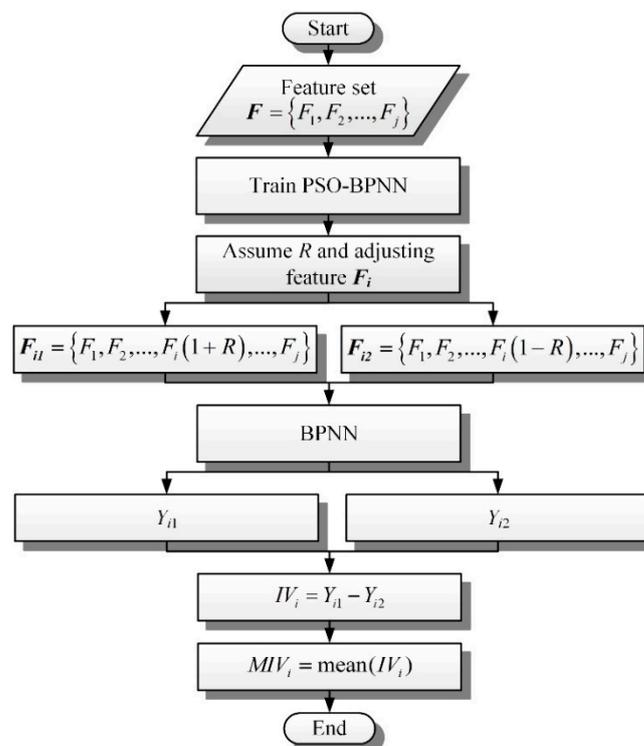


Figure 1. Flowchart of mean impact value (MIV).

3.2. Particle Swarm Optimization-BP Neural Network

Particle swarm optimization (PSO) is used to optimize the initial weights and bias of BPNN in this paper. This algorithm is a mathematical model developed by graphically representing bird behavior. The principle is to use two concepts of bird foraging: determining the direction according to its own experience and determining the direction by referring to the experience of others. Through these two concepts, the particles are randomly distributed in space. When the particles move, they refer to their best direction and the best direction of the group to determine the direction of the last movement, and iteratively find the optimal solution of the group [28]. The process of PSO-BP is shown in Figure 2. The detailed steps are as follows:

- Step (1) Set number of particles i , number of iterations t , the maximum number of iterations t_{max} , the acceleration constants c_1 and c_2 , and the inertia weights w .
- Step (2) Assume that the coordinates of each particle in space $X_i = (X_{1i}, X_{2i}, \dots, X_{Di})$, and the speed of each particle in space $V_i = (V_{1i}, V_{2i}, \dots, V_{Di})$.
- Step (3) Calculate the fitness values of all particles by BPNN, and obtain the best solution P_{best} for individuals and the best solution G_{best} for groups.
- Step (4) Correct flight speed of the particle $V_{i_new} = wV_i + c_1r_1(P_{best} - X_i) + c_2r_2(G_{best} - X_i)$
- Step (5) Correct the particle position $X_{i_new} = X_i + V_{i_new}$.
- Step (6) if $t < t_{max}$, $t = t + 1$ and repeat step (3) to step (5).
- Step (7) Obtain the best position of the groups as the best solution.
- Step (8) Get PSO-BP model.

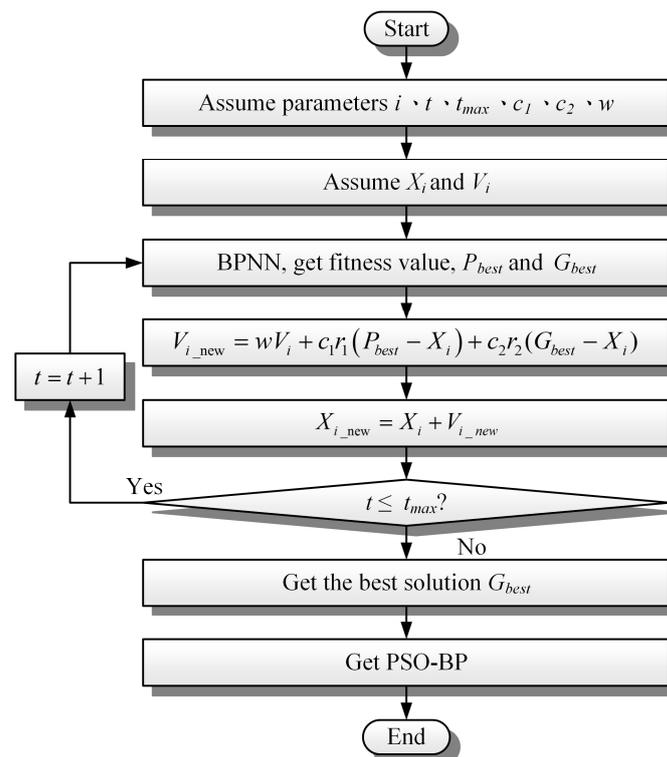


Figure 2. Flowchart of particle swarm optimization-back propagation (PSO-BP).

3.3. Feature Selection

In this study, the analysis result of the fault motor current signal contains 50 candidate features. According to the research in [14], it can be known that MIV reflects the influence of each feature on the classifier. Based on MIV, determining how to delete unimportant features is the key. In this study, MIV is combined with the PFS proposed in [22] to obtain a feature filtering process based on the MIV. The features corresponding to the smaller MIV are preferentially removed, and new feature vectors are established from the

remaining features. The goal is to optimize the initial weight and bias of BPNN, estimate the accuracy and ensure that it does not reduce the accuracy. Repeat the above process to delete unimportant features. The process of feature selection based on MIV is shown in Figure 3. The detailed steps are as follows:

- Step (1) Select all features to establish feature vector F_{origin} and set dimension $D = 50$.
- Step (2) Use PSO to optimize the initial weights w_{origin} and bias b_{origin} of BPNN.
- Step (3) Record the optimized result and evaluate the accuracy Acc_{remove} of BPNN.
- Step (4) Calculate MIV of all features.
- Step (5) Arrange MIV from the minimum to the maximum and remove a feature corresponding to the smallest MIV. $D = D - 1$.
- Step (6) Select unremoved features to create a new feature vector F_{new} , and evaluate the accuracy Acc_{remove} . If $Acc_{remove} > Acc_{origin}$, go back to Step (5).
- Step (7) Select the features that have not been removed to create a new feature vector F_{new} , and use PSO to optimize the initial weight w_{new} and b_{new} bias of BPNN. To evaluate the accuracy Acc_{new} .
- Step (8) If $Acc_{new} > Acc_{origin}$, go back to Step (5).

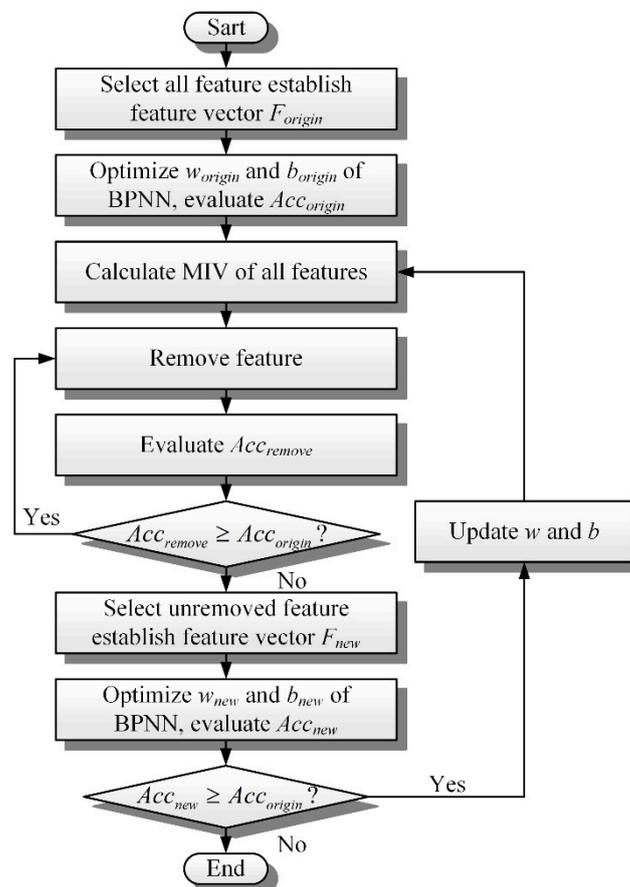


Figure 3. Flowchart of feature selection based on mean impact value (MIV).

4. Experimental Measurements and Analysis of IM

This section introduces the experimental equipment, experimental architecture, motor samples, and actual analysis methods of this study. After measuring the motors in four different conditions, using S-transform to analyze the current data. The motor characteristic curves of the four conditions are discussed in this section and introduce the four conditions of the motor and its characteristic curve in sequence.

4.1. Experiment Device

In this study, a four-pole IM (60 Hz/1.5 kW/1764 rpm) is used to drive a power platform composed of a servo motor (69 Hz/11 kW/2000 rpm) and a torque sensor. The data acquisition system (NI PXI-1033) captures the current signal and stores the data in a personal computer. The power meter platform, power meter platform control panel and data acquisition system are shown in Figure 4a–c:

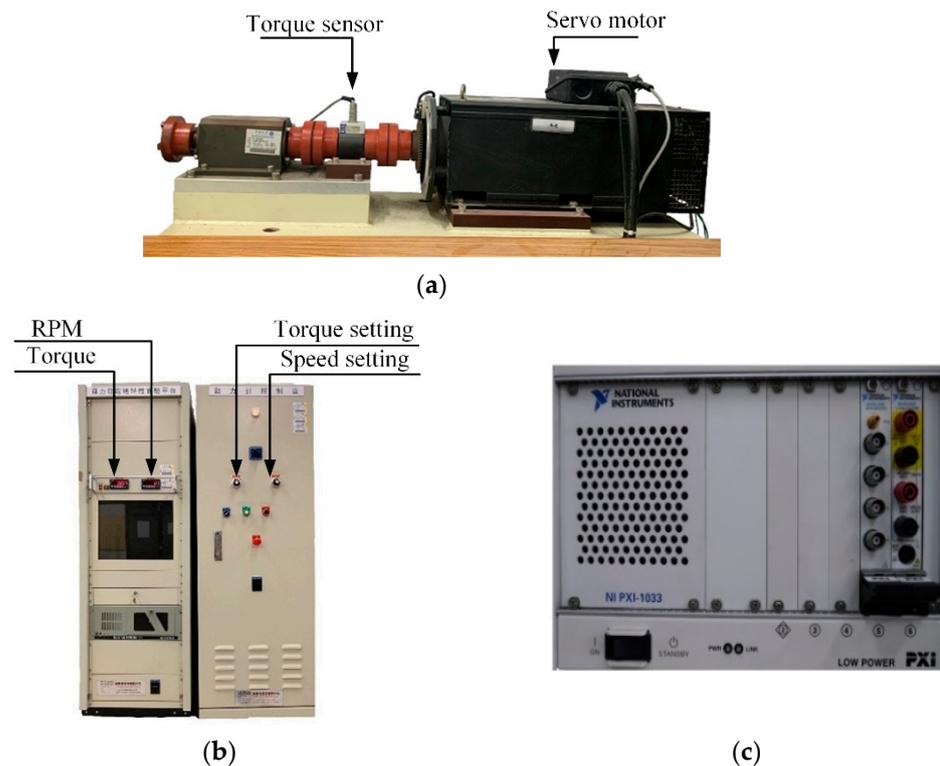


Figure 4. Experiment apparatus (a) power meter platform, (b) power meter platform control panel, and (c) data acquisition system.

4.2. Experiment Structure

In this experiment, the IM is used to drive the servo motor of the power meter platform, and the power platform generates the torque opposite to the IM as the load. The motor current signal of one phase is captured by the NI PXI-1033. The signal acquisition frequency is 1000 Hz, and each measurement time is 100 s. The healthy motor current signal, bearing damage motor current signal, stator layer short circuit current signal, and rotor drilling current signal are measured 200 s separately. The current signal is used to train the neural network and test the neural network. After S-transform, feature extraction is performed and input to a neural network for training and testing, and finally obtains the recognition rate. The schematic diagram of the experimental architecture is shown in Figure 5.

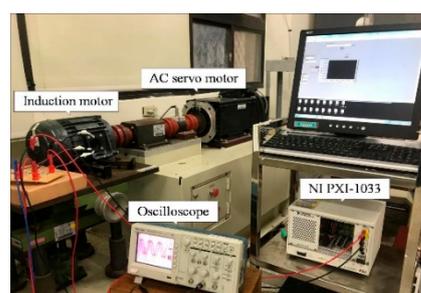


Figure 5. Experimental acquisition chain of the current signal.

4.3. Analysis Current of IM

According to IEEE-IAS, the types of faulty motors are bearings (44%), windings (26%), and rotors (8%). This study analyzes these types of faulty motors, and selects four types of motors: normal motors, bearing failure motors, stator short circuit fault motors, and rotor drilling fault motors.

This study analyzes the motor current signal through S-transform to obtain the ST matrix. From this, the ST matrix can obtain more characteristic curves in the time domain and the frequency domain. The steps show as follows: the maximum, minimum, average, mean square deviation, and standard deviation are extracted from each column of the ST matrix to establish 5 characteristic curves in time domain. The characteristic curves are: (1) Tmax; (2) Tmin; (3) Tmean; (4) Tmse; (5) Tstd; the maximum, minimum, average, mean square deviation and standard deviation are extracted from each row of the ST matrix to establish 5 characteristic curves in frequency domain. The characteristic curves are: (1) Fmax; (2) Fmin; (3) Fmean; (4) Fmse; (5) Fstd. In this section, the analysis results of the four status motors will be collated.

4.3.1. Healthy Motor

After the S-transform analysis of current signal of the healthy motor, the spectrum is shown in Figure 6a. The maximum amplitude is 0.3441, and the frequency is 121 Hz. The time domain characteristic curves Tmax, Tmin, Tmean, Tmse, and Tstd are shown in Figure 6b; the frequency domain characteristic curves Fmax, Fmin, Fmean, Fmse, and Fstd are shown in Figure 6c.

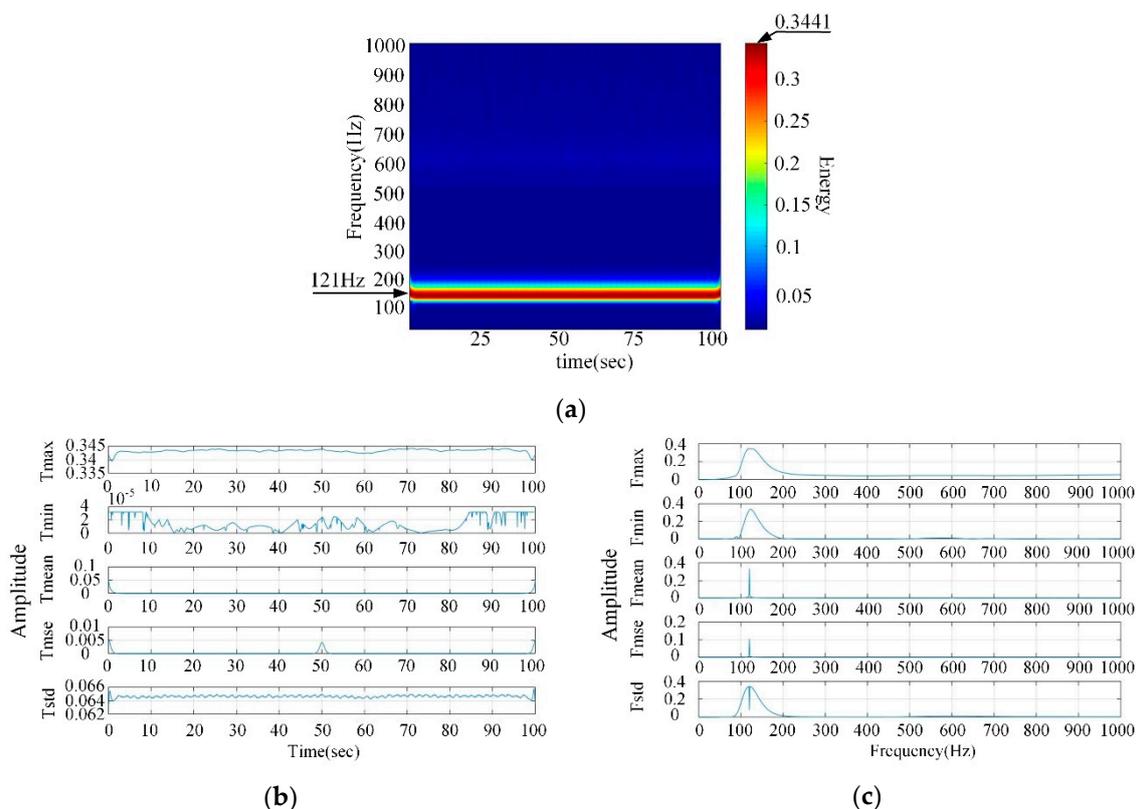


Figure 6. Healthy motor characteristic diagram (a) spectrum, (b) time domain characteristic curves, (c) frequency domain characteristic curves.

4.3.2. Bearing Failure Motor

After the S-transform analysis of current signal of the bearing failure motor, the spectrum is shown in Figure 7a. The maximum amplitude is 0.3312, and the frequency is 121 Hz. The time domain characteristic curves T_{max}, T_{min}, T_{mean}, T_{mse}, and T_{std} are shown in Figure 7b; the frequency domain characteristic curves F_{max}, F_{min}, F_{mean}, F_{mse}, and F_{std} are shown in Figure 7c. It can be found that the T_{max} characteristic curve of the bearing failure fault motor is mainly distributed at 0.33, which is slightly different from the T_{max} distribution of the healthy motor; most of the values on the T_{min} characteristic curve are close to 0; in the frequency domain characteristic curve, it is not much different from the healthy motor.

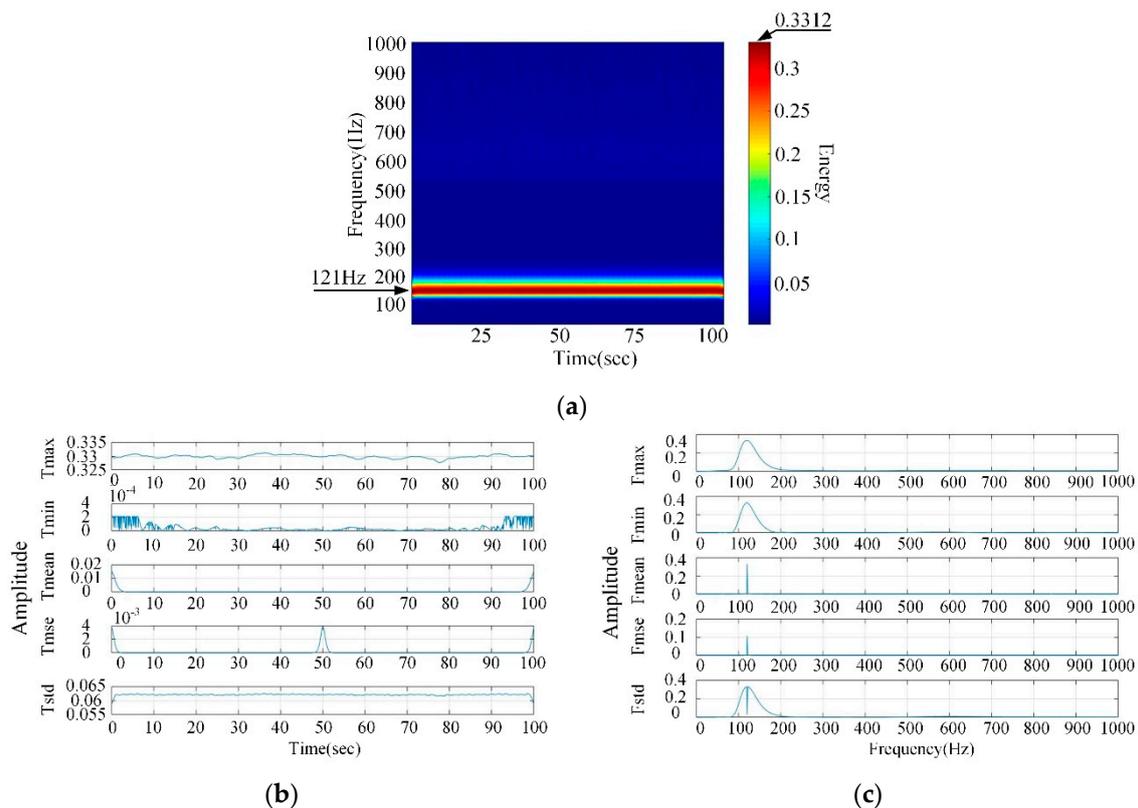


Figure 7. Bearing failure motor characteristic diagram (a) spectrum, (b) time domain characteristic curves, (c) frequency domain characteristic curves.

4.3.3. Stator Short Circuit Fault Motor

After S-transform analysis of current signal of the stator short circuit fault motor, the spectrum is shown in Figure 8a. The maximum amplitude is 0.3441, and the frequency is 121 Hz. The time domain characteristic curves T_{max}, T_{min}, T_{mean}, T_{mse}, and T_{std} are shown in Figure 8b; the frequency domain characteristic curves F_{max}, F_{min}, F_{mean}, F_{mse}, and F_{std} are shown in Figure 8c. It can be found that stator short circuit fault motor is less stable on the T_{max} characteristic curve; in the frequency domain characteristic curve, it is not much different from the healthy motor.

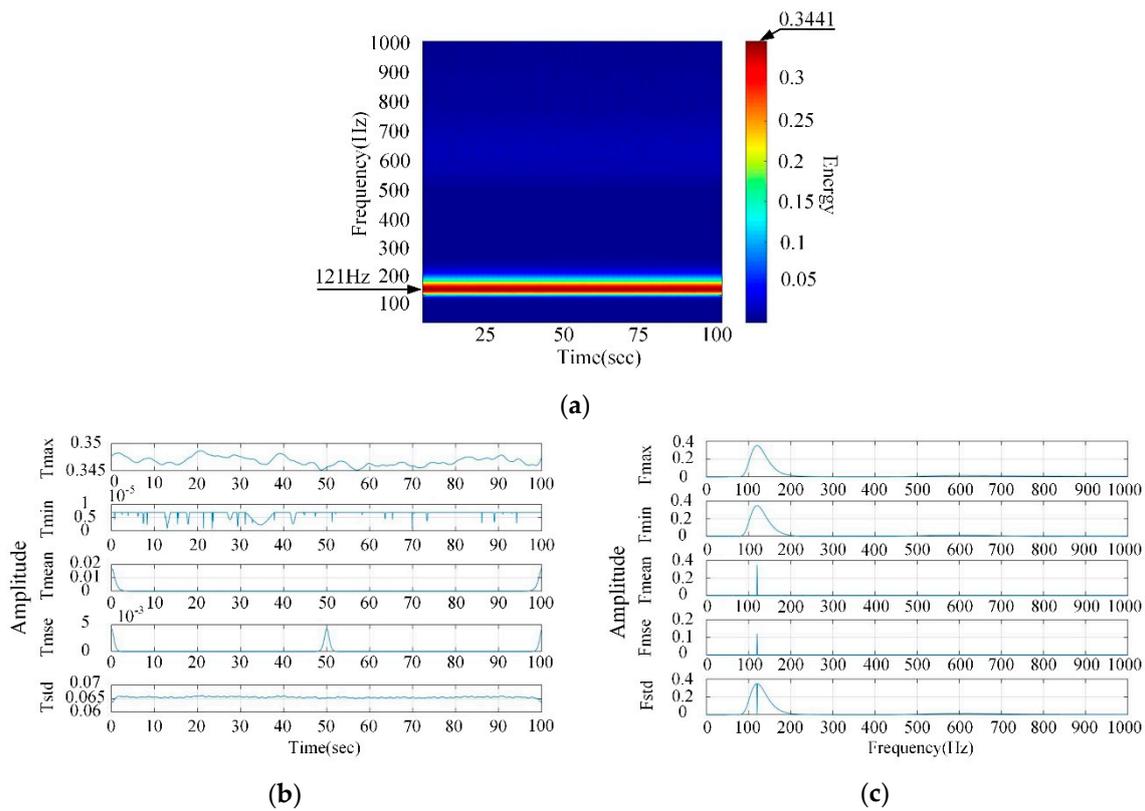


Figure 8. Stator short circuit fault motor characteristic diagram (a) spectrum, (b) time domain characteristic curves, (c) frequency domain characteristic curves.

4.3.4. Rotor Drilling Fault Motor

After S-transform analysis of current signal of the rotor drilling fault motor, the spectrum is shown in Figure 9a. The maximum amplitude is 0.3436, and the frequency is 121 Hz. The time domain characteristic curves T_{max} , T_{min} , T_{mean} , T_{mse} , and T_{std} are shown in Figure 9b; the frequency domain characteristic curves F_{max} , F_{min} , F_{mean} , F_{mse} , and F_{std} are shown in Figure 9c. It can be found that the rotor drilling fault motor is less stable on the T_{max} characteristic curve; most of the values on the T_{min} characteristic curve are close to 0; it is less stable on the T_{std} characteristic curve; in the frequency domain characteristic curve, it is not much different from the healthy motor.

4.4. Feature Extraction

The 10 characteristic curves obtained from the ST matrix from the time domain and the frequency domain in this study are as follows:

- (1) **T_{max}** : maximum value of each column of ST matrix.
- (2) **T_{min}** : minimum value of each column of ST matrix.
- (3) **T_{mean}** : average value of each column of ST matrix.
- (4) **T_{mse}** : mean square error of each column of ST matrix.
- (5) **T_{std}** : standard deviation of each column of ST matrix.
- (6) **F_{max}** : maximum value of each row of ST matrix.
- (7) **F_{min}** : minimum value of each row of ST matrix.
- (8) **F_{mean}** : average value of each row of ST matrix.
- (9) **F_{mse}** : mean square error of each row of ST matrix.
- (10) **F_{std}** : standard deviation of each row of ST matrix.

Then, the features are extracted from these 10 characteristic curves: (1) the sum of the maximum value and the minimum value; (2) the difference between the maximum value

and the minimum value; (3) the average value; (4) the mean square error; (5) the standard deviation, as shown in Table 1.

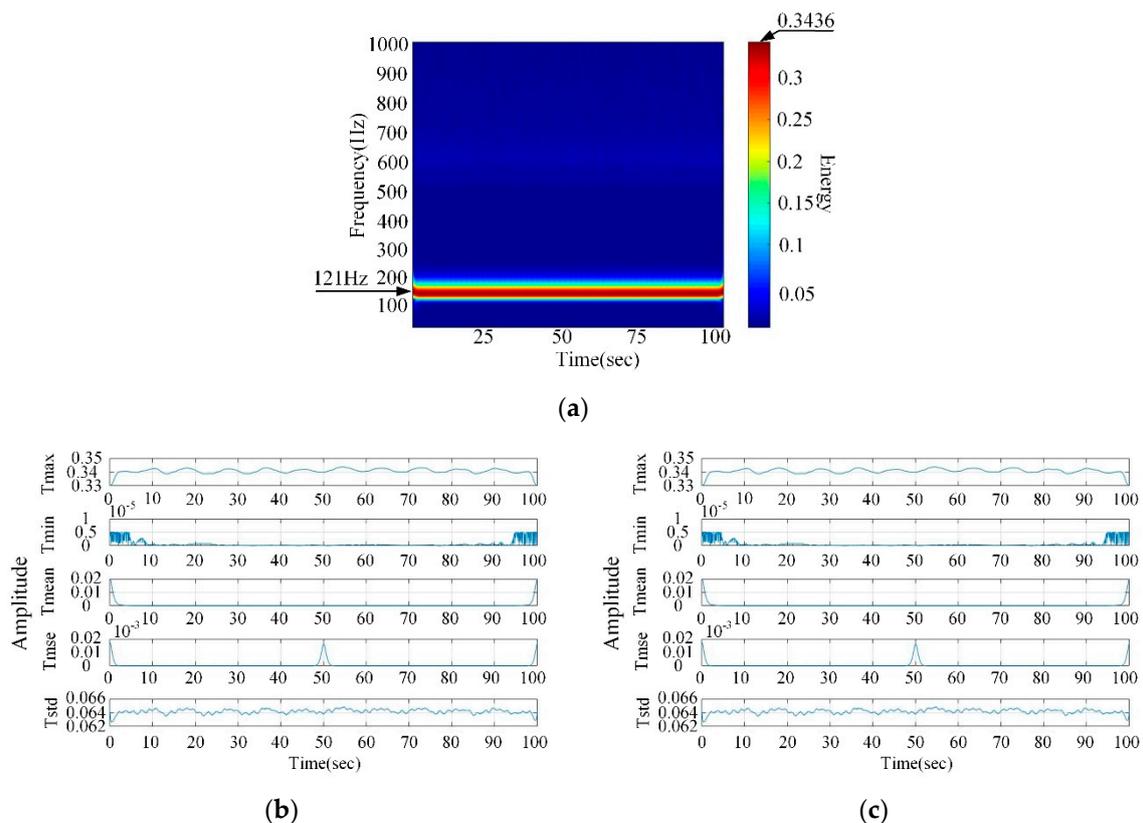


Figure 9. Stator short circuit fault motor characteristic diagram (a) spectrum, (b) time domain characteristic curves, (c) frequency domain characteristic curves.

Then, the features are extracted from these 10 characteristic curves: (1) the sum of the maximum value and the minimum value; (2) the difference between the maximum value and the minimum value; (3) the average value; (4) the mean square error; (5) the standard deviation, as shown in Table 1.

Table 1. Feature extraction.

	Tmax	Tmin	Tmean	Tmse	Tstd	Fmax	Fmin	Fmean	Fmse	Fstd
MAX + MIN	F1	F6	F11	F16	F21	F26	F31	F36	F41	F46
MAX – MIN	F2	F7	F12	F17	F22	F27	F32	F37	F42	F47
Mean	F3	F8	F13	F18	F23	F28	F33	F38	F43	F48
Mse	F4	F9	F14	F19	F24	F29	F34	F39	F44	F49
Std	F5	F10	F15	F20	F25	F30	F35	F40	F45	F50

In this study, a data acquisition system (NI PXI-1033) is used to record the IM current signal, and record 100 data for each fault condition, then the MATLAB compiler is used for S-transform analysis. Features are extracted from the analyzed data. Feature extraction is shown in Table 1. Then, the features are normalized. Therefore, the value is between 0 and 1, improving the efficiency of neural network training. The sample number is an even number selected as training data, and the sample number is an odd number selected as the test data, and input into BPNN for training and testing. Finally, the recognition success rate is calculated. The features of the time domain and the frequency domain after

S-transform are extracted according to Table 1, which are 50 characteristic values of F1, F2, F3, . . . F50. The process of feature extraction is shown in Figure 10. The feature distribution is shown in Figure 11, where the horizontal axis is the sample number, and the vertical axis is the feature number. It can be observed from Figure 11 that there are differences in the feature distribution of different fault conditions after the IM current signal analyzed by S-transform.

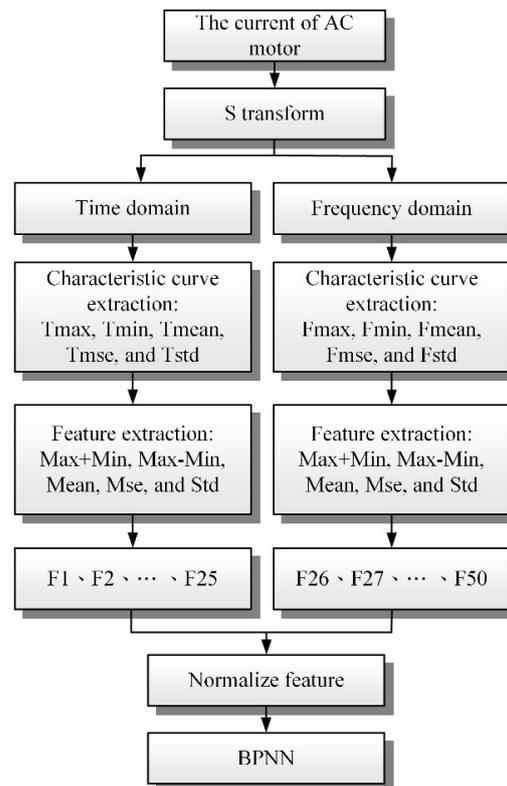


Figure 10. The process of feature extraction.

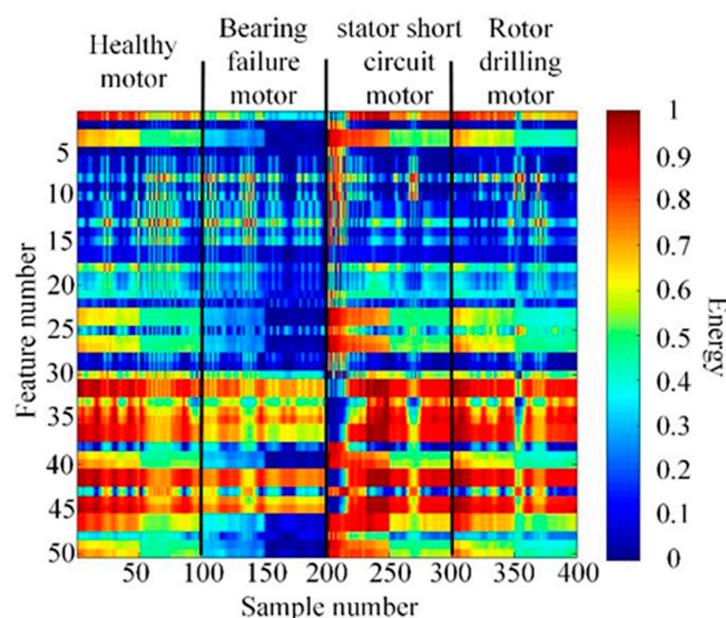


Figure 11. The distribution of features.

4.5. Classifier

In this study, it can be seen from the results of [12] that BPNN has a better effect on fault identification, so BPNN is chosen as classifier. We test the classification effect of different classifiers on noise-free current signals. It shows the accuracy of BPNN and probabilistic neural network (PNN) in Table 2. It can be seen that the accuracy of BPNN is better than PNN 10%. We use PSO to optimize the initial weights and bias of BPNN. We can see the classifier, PSO-BPNN, is the best model in of the three. As such, we use this model as a classifier. There are three layers of neurons in BPNN that we establish in this paper. The three layers are the input layer, hidden layer, and output layer. The number of neurons in the input layer is determined based on the identification features. The number of neurons in the hidden layer is 50. The number of neurons in the output layer is 4. These 4 neurons represent 4 different states of motors, healthy motor, bearing failure motor, stator short circuit fault motor, and rotor drilling fault motor. We conduct a multiclass prediction, and the ratio of training to testing data is 1:1. We measured 100 samples of motor current signals. From the 100 samples, it is cut according to the 1:1 ratio of training and testing.

Table 2. Accuracy of different classifier.

Classifier	Accuracy
PNN	86.12%
BPNN	96.3%
PSO-BPNN	100%

5. Results

The types of motors discussed in this study are normal motors, bearing failure motors, interlayer short circuit fault motors, and rotor drilling fault motors. In this section, the feature selection method in this study is used to select a new feature vector from the 50 identified features, and input to BPNN. We also test the effect of this method of feature selection on the accuracy in different noise ratios.

5.1. Motor Current Signal Measurement

This research measures the current signals of motors in various conditions according to the experimental structure in section IV. The signal sampling frequency is 1000 Hz, the sampling time is 100 s, a total of 100,000 sample points, and the captured signals are divided into 50 periods, and each period is 2000 sample points, each period is regarded as a sample of data. Each situation motor has a total of 50 training samples and 50 test samples. Using S-transform to analyze these samples to obtain the time characteristic curve and frequency characteristic curve, and then extract 50 identification features to normalize to establish a database of identification features. Considering that the current signals recorded in different measurement environments are affected by noise, this paper adds white gaussian noise (WGN), where the signal-to-noise ratio (SNR) is equal to 30 dB and 20 dB, Finally, a total of three kinds of databases are established, which are noise-free, SNR = 30 dB, and SNR = 20 dB.

5.2. Feature Selection Results

In this study, feature selection is performed on the features extracted after S-transform. The feature selection results of the three databases are shown in Tables 3–5. The selection process is use PFS combined with MIV. The noise-free database screening process is divided into four parts: feature number 50, feature number 16, feature number 10, and feature number 9, and the nine features after filtering are: F24, F31, F32, F24, F31, F32, F40, F43, F46, F47, F49, F50, the total recognition rate result is 99.4%. The result of BPNN computing time is 19.07 s, as shown in Table 3. The database selection process with SNR = 30 dB is divided into 4 parts: feature number 50, feature number 16, feature number 7, and feature number 6, after filtering the six features are: F38, F39, F40, F44, F46, F50, the total recognition rate

result is 86.2%, and the result of BPNN computing time is 18.90 s, as shown in Table 4. The database selection process with SNR = 20 dB is divided into four parts, namely feature number 50, feature number 27, feature number 3, and feature number 2, and the two features after filtering are: F42, F45. The total recognition rate result is 63.2%, the result of BPNN computing time is 18.76 s, as shown in Table 5. From Table 6, we can see that the selection features have a slight decrease in the identification effect of normal motors and short-circuit motors.

Table 3. Result of feature selection in the condition of signal-to-noise ratio (SNR) = ∞ .

Feature Number	Number of Features	Accuracy	Computing Time
F1, F2, F3, . . . , F48, F49, F50	50	100%	21.03
F11, F13, F24, F27, F31, F32, F34, F35, F39, F40, F42, F43, F46, F47, F49, F50	16	100%	19.43
F24, F31, F32, F39, F40, F43, F46, F47, F49, F50	10	100%	19.11
F24, F31, F32, F40, F43, F46, F47, F49, F50	9	99.4%	19.07

Table 4. Result of feature selection in the condition of SNR = 30 dB.

Feature Number	Number of Features	Accuracy	Computing Time
F1, F2, F3, . . . , F48, F49, F50	50	86.5%	21.09
F9, F20, F24, F26, F27, F30, F32, F36, F38, F39, F40, F44, F45, F46, F47, F50	16	88.6%	19.43
F36, F38, F39, F40, F44, F46, F50	7	86.5%	19.08
F38, F39, F40, F44, F46, F50	6	86.2%	18.90

Table 5. Result of feature selection in the condition of SNR = 20 dB.

Feature Number	Number of Features	Accuracy	Computing Time
F1, F2, F3, . . . , F48, F49, F50	50	64.0%	21.07
F1, F3, F5, F7, F8, F9, F11, F12, F13, F14, F20, F23, F24, F25, F26, F28, F32, F35, F36, F37, F38, F39, F42, F43, F45, F49, F50	27	71%	19.95
F42, F43, F45	3	67.3%	18.71
F42, F45	2	63.2%	18.76

Table 6. Motor identification rate under different fault conditions in SNR = ∞ .

Feature Selection Methods	Number of Features	Accuracy of Healthy Motor	Accuracy of Bearing Failure Motor	Accuracy of Stator Short Circuit Fault Motor	Accuracy of Rotor Drilling Fault Motor	Total Accuracy
All features	50	100%	100%	100%	100%	100%
MIV base on PSO-BPNN	9	97.7%	100%	99.9%	100%	99.4%

We can see the process of feature selection from Tables 3–5. During the screening process, the recognition rate will increase. Under SNR = 30 dB, the recognition rate will increase from 86.5% to 88.6% during the feature screening process; under the condition of 20 dB, the recognition rate is increased from 64% to 71% during the feature selection process. In the selection result, the accuracy will be slightly reduced, but the feature will be the best combination. In the case of SNR = ∞ the accuracy is reduced from 100% to 99.4%, and the calculation time is reduced from 21.03 s to 19.07 s; under the condition of SNR = 30 dB, the

accuracy is reduced from 86.5% to 86.2%, and the calculation time is 21.09 s is reduced to 18.09 s; under the condition of SNR = 20 dB, the accuracy is reduced from 64% to 63.2%, and the calculation time is reduced from 21.07 s to 18.76 s. From these results, it can be seen that the feature selection used in this paper can reduce most of the recognition features and reduce the BPNN computing time while maintaining the recognition rate. We tested the effect of different selection methods on noise-free current signal. The result show in Table 7. We selected two feature selection methods, genetic algorithm (GA) and reliefF, for comparison. It can be seen that the method proposed in this paper has the same as relief in terms of the number of features. MIV base on PSO-BPNN has 19 fewer features than GA selection method in feature number. In terms of recognition rate, there is still a 99.4% recognition success rate. It can see that the feature selection method proposed in this paper is better than the other two. Table 8 shows the confusion matrix of classification result in SNR = ∞ . This table is the average of 20 executions. The number of test data is 50. In 20 running times, a total of 2 healthy motor signals are classified as stator short circuit fault motor signals. A total of 4 healthy motor signals are classified as rotor drilling fault motor signals. A total of 1 stator short circuit fault motor signal is classified as healthy.

Table 7. Results of different feature selection methods.

Feature Selection Methods	Number of Features	Accuracy
All features	50	100%
GA	28	100%
ReliefF	9	98.5%
MIV base on PSO-BPNN	9	99.4%

Table 8. Confusion matrix of classification result in SNR = ∞ .

	Healthy Motor	Bearing Failure Motor	Stator Short Circuit Fault Motor	Rotor Drilling Fault Motor	Total
Healthy motor	48.85	0	0.05	0	99.8% 0.2%
Bearing failure motor	0	50	0	0	100% 0%
Stator short circuit fault motor	0.1	0	49.9	0	100% 0%
Rotor drilling fault motor	0.2	0	0	50	97.8% 2.2%
Total	97.7% 2.3%	100% 0%	99.8% 0.2%	100% 0%	99.4% 0.6%

6. Conclusions

This paper discusses the identification and classification of IM bearing damage, stator interlayer short circuit, and rotor drilling. It includes the analysis method, the calculation MIV of each feature, the optimization of the classification model, and the combination of PFS and MIV in features selection can achieve a certain effect. After the IM current signal is converted by S-transform, a total of 50 features are extracted from these data. After the selection process of PFS combined with MIV, 50 features can be reduced to 9 features, and the accuracy can reach 99.4%, which is almost the same as the recognition rate of other large numbers of features. Therefore, this method can not only reduce the number of features, but also achieve a similar recognition rate; in terms of computing time, it is also reduced from 21.03 s to 19.07 s, reducing the computing time by 9%. The 9 features after selection are better than 50 features under the comprehensive consideration of recognition rate and

computing time. It is shown that the method of filtering features is helpful to classify IM faults in BPNN.

Author Contributions: Conceptualization, C.-Y.L. and H.-Y.O.; Methodology, C.-Y.L. and H.-Y.O.; Software, C.-Y.L. and H.-Y.O.; Validation, C.-Y.L. and H.-Y.O.; Formal Analysis, C.-Y.L. and H.-Y.O.; Investigation, C.-Y.L. and H.-Y.O.; Resources, C.-Y.L. and H.-Y.O.; Data Curation, C.-Y.L. and H.-Y.O.; Writing-Original Draft Preparation, C.-Y.L. and H.-Y.O.; Writing-Review & Editing, C.-Y.L. and H.-Y.O.; Visualization, C.-Y.L. and H.-Y.O.; Supervision, C.-Y.L.; Project Administration, C.-Y.L.; Funding Acquisition, C.-Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

- Brkovic, A.; Gajic, D.; Gligorijevic, J.; Savic Gajic, I.; Georgieva, O.; Gennaro, S.-D. Early fault detection and diagnosis in bearings for more efficient operation of rotating machinery. *Energy* **2017**, *136*, 63–71. [[CrossRef](#)]
- Ro, H.S.; Kim, D.H.; Jeong, H.G.; Lee, K.-B. Tolerant control for power transistor faults in switched reluctance motor drives. *IEEE Trans. Ind. Appl.* **2015**, *51*, 3187–3197. [[CrossRef](#)]
- Tobon-Mejia, D.A.; Medjaher, K.; Zerhouni, N. CNC machine tool's wear diagnostic and prognostic by using dynamic Bayesian networks. *Mech. Syst. Signal Process.* **2012**, *28*, 167–182. [[CrossRef](#)]
- Resendiz-Ochoa, E.; Osornio-Rios, R.A.; Benitez-Rangel, J.P.; De Romero-Troncoso, R.J.; Morales-Hernandez, L.A. Induction motor failure analysis: An automatic methodology based on infrared imaging. *IEEE Access* **2018**, *6*, 76993–77003. [[CrossRef](#)]
- Hsueh, Y.-M.; Ittangihal, V.R.; Wu, W.-B.; Chang, H.-C.; Kuo, C.-C. Fault Diagnosis System for Induction Motors by CNN Using Empirical Wavelet Transform. *Symmetry* **2019**, *11*, 1212. [[CrossRef](#)]
- Korkua1, S.; Jain, H.; Lee, W.-J.; Kwan, C. Wireless Health Monitoring System for Vibration Detection of Induction Motors. In Proceedings of the 2010 IEEE Industrial and Commercial Power Systems Technical Conference, Tallahassee, FL, USA, 9–13 May 2010.
- Ge, M.; Wand, J.; Xu, Y.; Zhang, F.; Bai, K.; Ren, X. Rolling Bearing Fault Diagnosis Based on EWT Sub-Modal Hypothesis Test and Ambiguity Correlation Classification. *Symmetry* **2018**, *10*, 730. [[CrossRef](#)]
- Yang, T.; Pen, H.; Wang, Z.; Chang, C.S. Feature Knowledge Based Fault Detection of Induction Motors through the Analysis of Stator Current Data. *IEEE Trans. Instrum. Meas.* **2016**, *65*, 549–558. [[CrossRef](#)]
- Oyamada, M. Extracting feature engineering knowledge from data science notebooks. In Proceedings of the IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019.
- Rauber, T.W.; de Boldt, F.A.; Varejão, F.M. Heterogeneous feature models and feature selection applied to bearing fault diagnosis. *IEEE Trans. Ind. Electron.* **2014**, *62*, 637–646. [[CrossRef](#)]
- Gligorijevic, J.; Gajic, D.; Brkovic, A.; Savic Gajic, I.; Georgieva, O.; Gennaro, S.-D. Online Condition Monitoring of Bearings to Support Total Productive Maintenance in the Packaging Materials Industry. *Sensors* **2016**, *16*, 316. [[CrossRef](#)]
- Huang, Z.-L.; Zhang, J.; Zhao, T.-H.; Sun, Y. Synchrosqueezing S-transform and its application in seismic spectral decomposition. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 817–825. [[CrossRef](#)]
- Zhang, S.; Huang, D. EEG time-frequency analysis based on the improved S-transform. In Proceedings of the 25th Chinese Control and Decision Conference, Guiyang, China, 25–27 May 2013.
- Feng, Z.P.; Chen, X.W.; Liang, M. Iterative generalized synchrosqueezing transform for fault diagnosis of wind turbine planetary gearbox under nonstationary conditions. *Mech. Syst. Signal Process.* **2015**, *52*, 360–375. [[CrossRef](#)]
- Wu, H.T.; Chan, Y.H.; Lin, Y.T.; Yeh, Y.H. Using synchrosqueezing transform to discover breathing dynamics from ECG signals. *Appl. Comput. Harmon. Anal.* **2014**, *36*, 354–359. [[CrossRef](#)]
- Wang, Z.-C.; Ren, W.-X.; Liu, J.-L. A synchrosqueezed wavelet transform enhanced by extended analytical mode decomposition method for dynamic signal reconstruction. *J. Sound Vib.* **2013**, *332*, 6016–6028. [[CrossRef](#)]
- Wang, S.; Chen, X.; Wang, Y.; Cai, G.; Ding, B.; Zhang, X. Nonlinear squeezing time-frequency transform for weak signal detection. *Signal Process.* **2015**, *113*, 195–210. [[CrossRef](#)]
- Ahrabian, A.; Looney, D.; Stanković, L.; Mandić, D.P. Synchrosqueezing-based time-frequency analysis of multivariate data. *Signal Process.* **2015**, *106*, 331–341. [[CrossRef](#)]
- Kwan, C.; Zhang, X.; Xu, R.; Haynes, L. A novel approach to fault diagnostics and prognostics. In Proceedings of the 2003 IEEE International Conference on Robotics and Automation, Taipei, Taiwan, 14–19 September 2003; pp. 604–609.

20. Yang, C.; Lu, L.; Lin, H.; Guan, R.; Shi, X.; Liang, Y. A Fuzzy-Statistics-Based Principal Component Analysis (FS-PCA) Method for Multispectral Image Enhancement and Display. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 3937–3947. [[CrossRef](#)]
21. Zhang, Y.; Wang, Z. A hybrid model for blood pressure prediction from a PPG signal based on MIV and GA-BP neural network. In Proceedings of the 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, Guilin, China, 29–31 July 2017.
22. Lee, C.-Y.; Shen, Y.-X. Optimal feature selection for power-quality disturbances classification. *IEEE Trans. Power Deliv.* **2006**, *26*, 2342–2351. [[CrossRef](#)]
23. Qian, T.; Li, X.; Ayhan, B.; Xu, R.; Kwan, C.; Griffin, T. Application of Support Vector Machines to Vapor Detection and Classification for Environmental Monitoring of Spacecraft. In *ISNN 2006. Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2006.
24. Liu, N.; Gao, J.; Zhang, B.; Li, F.; Wang, Q. Time-Frequency Analysis of Seismic Data Using a Three Parameters S Transform. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 142–146. [[CrossRef](#)]
25. Brown, R.A.; Lauzon, M.L.; Frayne, R. A general description of linear time–frequency transforms and formulation of a fast, invertible transform that samples the continuous S-transform spectrum nonredundantly. *IEEE Trans. Signal Process.* **2010**, *58*, 281–290. [[CrossRef](#)]
26. Kohavi, R.; John, G.H. Wrappers for feature subset selection. *Artif. Intell.* **1997**, *97*, 273–324. [[CrossRef](#)]
27. Liu, X.; Tang, J. Mass classification in mammograms using selected geometry and texture features and a new SVM-based feature selection method. *IEEE Syst. J.* **2014**, *8*, 910–920. [[CrossRef](#)]
28. Fernandez-Martinez, J.L.; Garcia-Gonzalo, E. Stochastic Stability Analysis of the Linear Continuous and Discrete PSO Models. *IEEE Trans. Evol. Comput.* **2011**, *15*, 405–423. [[CrossRef](#)]