

Article

Feature Ranking and Differential Evolution for Feature Selection in Brushless DC Motor Fault Diagnosis

Chun-Yao Lee * and Chen-Hsu Hung

Department of Electrical Engineering, Chung Yuan Christian University, No. 200, Zhongbei Road, Zhongli District, Taoyuan City 320, Taiwan; g10978021@cycu.edu.tw

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

Abstract: A fault diagnosis system with the ability to recognize many different faults obviously has a certain complexity. Therefore, improving the performance of similar systems has attracted much research interest. This article proposes a system of feature ranking and differential evolution for feature selection in BLDC fault diagnosis. First, this study used the Hilbert–Huang transform (HHT) to extract the features of four different types of brushless DC motor Hall signal. When there is a fault, the symmetry of the Hall signal will be influenced. Second, we used feature selection based on a distance discriminant (FSDD) to calculate the feature factors which base on the category separability of features to select the features which have a positive correlation with the types. The features were entered sequentially into the two supervised classifiers: backpropagation neural network (BPNN) and linear discriminant analysis (LDA), and the identification results were then evaluated. The feature input for the classifier was derived from the FSDD, and then we optimized the feature rank using differential evolution (DE). Finally, the results were verified from the BLDC motor's operating environment simulation with the same features by adding appropriate signal-to-noise ratio magnitudes. The identification system obtained an accuracy rate of 96% when there were 14 features. Additionally, the experimental results show that the proposed system has a robust anti-noise ability, and the accuracy rate is 92.04%, even when 20 dB of white Gaussian noise is added to the signal. Moreover, compared with the systems established from the discrete wavelet transform (DWT) and a variety of classifiers, our proposed system has a higher accuracy with fewer features.

Keywords: fault detection; feature selection; backpropagation neural network



Citation: Lee, C.-Y.; Hung, C.-H. Feature Ranking and Differential Evolution for Feature Selection in Brushless DC Motor Fault Diagnosis. *Symmetry* **2021**, *13*, 1291. <https://doi.org/10.3390/sym13071291>

Academic Editor: Daniele Vilone

Received: 5 June 2021

Accepted: 15 July 2021

Published: 18 July 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 response to global environmental issues, environmental awareness and carbon emissions issues have received much attention. The demand for hybrid electric vehicles (HEVs) and electric vehicles (EVs) has gradually increased [1]. Attempts have been made to use HEVs and EVs instead of traditional vehicles, in order to reduce carbon emissions. The following topics need to be paid attention: the brushless DC motor (BLDC) that acts as a core piece of equipment in EVs [2], the range anxiety issue [3] and energy storage systems [4,5]. As time passes by, and with more research and development invested in EVs, such as sensitivity analysis of a rolling stock hydrogen hybrid powertrain [6], and a battery thermal management system [4,5], the work in this article constructed a system to identify the BLDC fault types. In order to detect the operation status of the BLDC, Hall sensors or sensorless algorithms based on back electromotive force are commonly used [7]. The three Hall sensors are installed into the BLDC motor with 120-degree phase differences, so the Hall signals are 120 electrical degree phase differences in normal. When there is a fault, the symmetry of the Hall signal will be affected. A Hall sensor has the obvious advantages of a low cost and simple structure [8]. Additionally, DC motors using Hall sensors have been widely used in commercial and industrial applications [9]. Therefore, this article uses the Hall signal, which is an electrical technology, to establish an identification system. The motor may suffer from different failures, including stator failure [10–12], rotor

failure [13,14], bearing failure [15,16], eccentricity fault [17] and inverter fault [18]. Stator failure accounts for 30% to 40% of the total failures in motors; rotor failure accounts for 5% to 10% of the total failures; and bearing failure accounts for 40% to 50% of total failures [19]. If the motor is running in a fault condition for a long time, it may cause operation economic losses due to poor performance, which may affect driving safety.

The comprehensive description of the proposed fault diagnosis system includes signal analysis, feature selection, optimization and classifiers. Signal analysis has been developed for decades, and the Hilbert–Huang transform (HHT) is based on the intrinsic mode functions (IMFs) of the original signal to calculate the instantaneous frequency, and then perform spectrum analysis [20]. It is necessary to calculate the IMF before fault diagnosis. When adding more IMFs, the calculation cost will be increased. If the IMF is insufficient, the representative information may be neglected [21]. However, since there is no need to choose the mother wavelet, it is not affected by the resolution of the time domain and frequency domain, and this can more accurately decompose the signal in the high-frequency domain. This technique is commonly used for fault detection in the fields of biomedical applications, structural testing and rotating machinery [22].

After signal analysis, the original signal can provide features with a good identification rate through feature selection. Feature selection can be divided into filter, wrapper, hybrid approach and embedded feature selection [23]. The filter type is based on the relationship between features as the criterion [24], and the wrapper type is based on the relationship between features and the target variable as the criterion [25]. The embedded type is usually used for high-dimensional data features [26,27]. The wrapper and hybrid approach types can obtain better results [28], but the filter type is usually used when considering the computational cost and a large number of features [26,27]. Therefore, this study used the filter feature selection to calculate feature weights, such as the FSDD, belonging to a clustering algorithm [29].

The FSDD calculates the distance discriminant factor of features that are based on the variance distance between feature clusters and within clusters. Due to the low computational complexity of the FSDD, it is suitable for high-dimensional problems or online feature selection [30]. The feature cannot obtain high numerical factors if they have good separation features, but there is a large distance within clusters. The features were obtained through signal analysis in this study, and then the factors of each feature were obtained from the FSDD. The features were entered into the classifier in the order of factors, and then the types of motor failures were identified by the classifier.

This article combines DE with feature factors after feature selection to optimize the feature ranking. DE is an effective and simple global optimization algorithm. The convergence speed and robustness of common benchmark functions and practical problems are better than those of many algorithms [31]. Finally, the recognition results of the BPNN were compared with the result of LDA. An artificial neural network (ANN) is a common nonlinear function processor that imitates the structure and pattern of the human brain [32]. The performance of the learning process of the neural network depends on the weights of the neural network in the training phase. A BPNN [33] is a supervised machine learning technique that adjusts its weights to minimize the error of the calculated output [33], and it is suitable for identifying nonlinear relationships [34]. In a feedforward ANN, the data flow has no feedback paths [35]. A BPNN is used in the fault diagnosis problem of NPC inverters [36], high-impedance faults [37] and virtual speed sensors for DC motors [38], while LDA is often used for supervised feature extraction and can also maximize the variance between classes based on linear projections, minimize the intra-class variance and finally obtain the largest separation between the feature sets in each class [39]. LDA can be used as a supervised classifier, through different types of data to maximize separation [40].

Based on the above-mentioned related literature, this research proposes a fault identification system for a BLDC established by Hall signals, which includes signal analysis selection, feature selection and classifiers.

2. Experimental Setup and Hardware Design

2.1. Experimental Design

This section introduces the experimental equipment, experimental architecture and signal samples in this research and studies the healthy and three different types of faults in a BLDC to build a fault classification system. A total of 750 samples of Hall signal data records for each motor and the measured signals were analyzed by the HHT in Matlab. After the analysis, the extracted features that can reflect the motor conditions were normalized so that the feature values of the 4 motor types were between 0 and 1, which avoids the gradient explosion problem in the classifier. Additionally, we will discuss the signals of the four conditions to grasp the operating conditions of the motor.

The BLDC (420 W/3020 rpm/DC 24 V/60 Hz) had the following three fault conditions in this experiment: bearing fault, stator winding fault and rotor fault. The data acquisition system (NI PXIe-1073) was used to acquire the Hall signal of the DC brushless motor, and the sampling rate was 1000 Hz, and the measurement time was 1500 s. There was a total of 1500 s of measurement records for BLDC motors in each condition, and the 1500-s data were divided into 750 samples of data, every sample having 2000 points. The load was simulated by an AC servo motor (11 kW/2000 rpm/69 Hz). Through the above equipment, the measurement data of the BLDC Hall signal can be obtained.

2.2. Experimental Architecture

The process of this research is that the servo motor of the dynamometer generates the opposite torque to the BLDC as the load, and then the BLDC motor drives the operation. The BLDCM rated voltage was 24 V, and the motor rate speed was configured as 3020 RPM. The BLDCM parameters are listed in Table 1. A total of four BLDCs were tested in this experiment. One motor was normal, whereas the other three motors were faulty. The faulty types included bearing damage in the inner raceway, winding short circuit and rotor damage. The bearing inner raceway had a 1 mm physical crack. The winding short circuit was set by exfoliating a part of the 2-coil insulation. The rotor damage was set by digging a hole. The NI PXIe-1073 was used to capture the Hall signal during operation. The data from the measurement were analyzed by Matlab for signal analysis, and then feature selection was used to calculate the factor of the features. The rank of features was optimized by DE after the features were ranked in descending order by the feature factors. Finally, the results of the fault type identification from classifiers were returned. The experimental process and configuration are shown in Figure 1.

Table 1. BLDC parameters.

Type	Rated Current	Rated Torque	Rated Speed	Rated Output Power	Rated Efficiency
BL5K35030D	22.07 A	13.5 Kg-cm	3020 RPM	418.7 W	81.2%

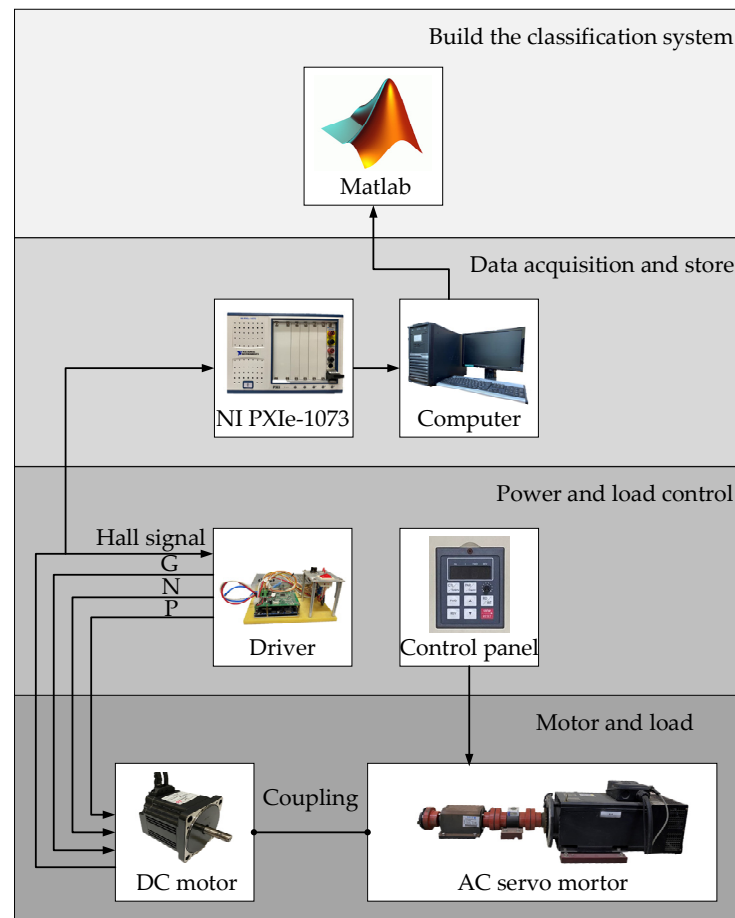


Figure 1. Equipment layout for capturing the signal of the BLDC.

3. Proposed Method

3.1. Signal Analysis and Feature Extraction

The signal can be expressed in the time and frequency domains. In some cases, the frequency domain of the signal can be presented in a clearer way than the time domain [41]. Additionally, the signals are usually significantly different elements. These scores cannot be expressed in the same base function, meaning two or more base functions are required to analyze the signal separately [42].

Dr. Norden E. Huang proposed the HHT in 1998, and it has since been widely used in speech analysis and nonlinear and unstable signal analysis [43]. Empirical mode decomposition is the basic theory of the HHT. IMFs through the Hilbert transform can be used to obtain the Hilbert spectrum of the analysis data.

The original function of the input is represented by $x(t)$, which can be decomposed into n IMFs and trend functions through EMD. EMD is modeled as

$$x(t) = \sum_{l=1}^n h_l(t) + r_n(t) \quad (1)$$

Then, IMFs are brought into the Hilbert transform (HT) to obtain the instantaneous amplitude and instantaneous frequency of the signals. The HT is modeled as

$$H_l(t) = \frac{1}{\pi} P_v \int_{-\infty}^{\infty} \frac{h_l(\tau)}{t - \tau} d\tau \quad (2)$$

where Pv is the caution principal value, which is to avoid the singularity of $\tau = t$ and $\tau = \pm\infty$. The Hilbert spectrum is formulated as

$$Z_l(t) = h_l(t) + jH_l(t) = a_l(t)e^{j\theta_l(t)} \quad (3)$$

where $h_l(t)$ represents the IMF, and $H_l(t)$ can be obtained through Hilbert transformation. Among them, $a_l(t)$ is the instantaneous amplitude, and $\theta_l(t)$ is the instantaneous phase angle.

The four types of motor Hall signals were decomposed through empirical mode decomposition, which separates the signal from the first to eighth layers (IMF1 to IMF8), and the instantaneous amplitude and instantaneous frequency of each layer were obtained through the Hilbert–Huang transform. Additionally, we then captured the maximum (Tmax), average (Tmean), mean square error (Tmse), standard deviation (Tstd), maximum/mean (Tmax/Tmean) and maximum/root mean square (Tmax/Trms) of the time domain, and the maximum (Fmax), average (Fmean), mean square error (Fmse), standard deviation (Fstd), maximum/average (Fmax/Fmean) and maximum/root mean square (Fmax/Trms) of the frequency domain. Each IMF took 12 features and normalized them so that the feature values of the 4 motor types were distributed between 0 and 1. This step obtained a total of 96 features, as shown in Table 2.

Table 2. Feature extraction of the HHT.

		Max	Mean	Mse	Std	Max/Mean	Max/Rms
Time domain	IMF1	F1	F2	F3	F4	F5	F6
	IMF2	F7	F8	F9	F10	F11	F12
	IMF3	F13	F14	F15	F16	F17	F18
	IMF4	F19	F20	F21	F22	F23	F24
	IMF5	F25	F26	F27	F28	F29	F30
	IMF6	F31	F32	F33	F34	F35	F36
	IMF7	F37	F38	F39	F40	F41	F42
	IMF8	F43	F44	F45	F46	F47	F48
Frequency domain	IMF1	F49	F50	F51	F52	F53	F54
	IMF2	F55	F56	F57	F58	F59	F60
	IMF3	F61	F62	F63	F64	F65	F66
	IMF4	F67	F68	F69	F70	F71	F72
	IMF5	F73	F74	F75	F76	F77	F78
	IMF6	F79	F80	F81	F82	F83	F84
	IMF7	F85	F86	F87	F88	F89	F90
	IMF8	F91	F92	F93	F94	F95	F96

The normal BLDC Hall signal was decomposed by the empirical mode, and it generated the IMF waveform, as shown in Figure 2a. The transform first extracted a signal of a high frequency, and the subsequent layers of the IMF were low-frequency waveforms.

The feature map is a schematic of the features from 3000 pieces of data, and the feature map was drawn by Matlab. The vertical axis of the feature map is the number of features, and the horizontal axis is the number of data. As it is shown in Figure 2b, the features from the four types of motors are highly similar, and there is no obvious difference.

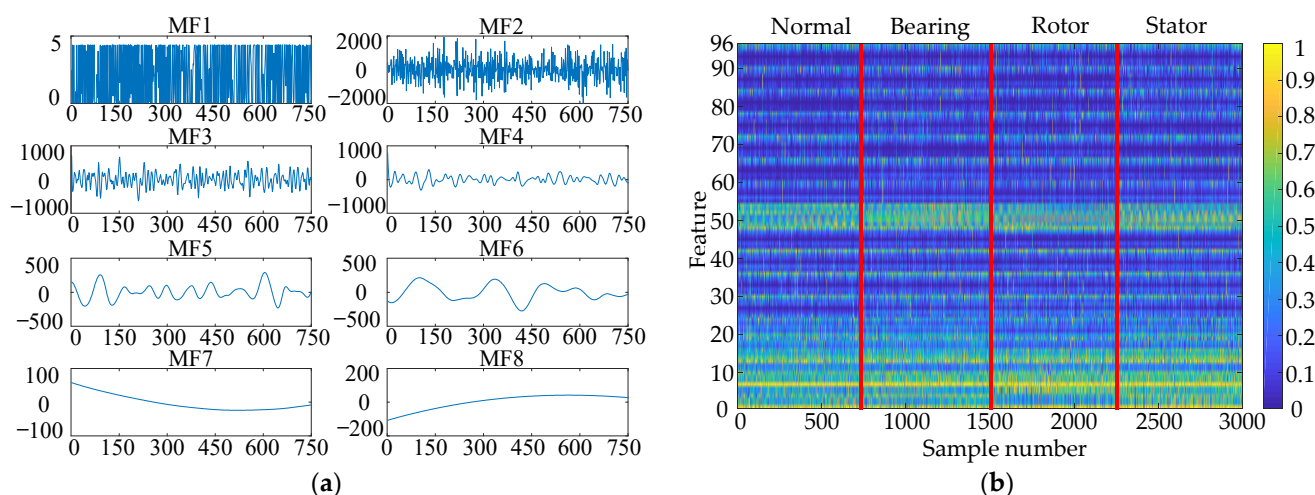


Figure 2. (a) IMF of normal motor; (b) feature distribution of the HHT.

3.2. Feature Selection

The system includes feature selection in order to pick out the few critical features from the signal. The feature selection was implemented using the FSDD of the clustering algorithm to calculate the category separability of features. The high value of the factor represents an important feature. Additionally, the features were ranked in descending order by feature factors after feature selection. The features of the Hall signal extracted after signal analysis can reduce the recognition rate of the classifier or not affect the recognition result by feature selection and deletion, which can save the cost of the calculation of the recognition system.

The feature distance discriminant factor λ^m is based on the Euclidean distance between the features of the same category d_w^m and the Euclidean distance between the features of different classes d_b^m . The flow chart of the FSDD is shown in Figure 3. The Euclidean distance of the feature was calculated by the center of the category feature g_c^m and the center of the sample feature g_i^m , where C , m and i are the category number, feature number and sample number. q_i^m is the feature of the sample. The compensation factor η^m was calculated by the distance variance u_b^m and v_w^m . The calculation procedure is as follows:

Step 1. Calculate the variance and average of all the samples in the m th feature.

$$\sigma_m^2 = \frac{1}{N} \sum_{i=1}^N (q_i^m - \bar{q}^m)^2 \quad (4)$$

$$\bar{q}^m = \frac{1}{N} \sum_{i=1}^N q_i^m \quad (5)$$

Step 2. Calculate the variance and the average of the sample of class C in the m th feature.

$$\sigma_{m(C)}^2 = \frac{1}{N_C} \sum_{i=1}^{N_C} (q_i^m - \bar{q}_C^m)^2 \quad (6)$$

$$\bar{q}_C^m = \frac{1}{N_C} \sum_{i=1}^{N_C} q_i^m \quad (7)$$

Step 3. Calculate the weighted variance of the class center g_C at the m th feature.

$$\sigma_m^2 = \sum_{C=1}^C \rho_C (g_C^m - g^m)^2 \quad (8)$$

$$g^m = \sum_{C=1}^C \rho_C g_C^m \quad (9)$$

Step 4. Calculate the inter-class distance d_b of the m th feature and the intra-class distance d_w of the m th feature.

$$d_b^m = \frac{\sigma_m^2}{\sigma_m^2} \quad (10)$$

$$d_w^m = 2 \frac{\sum_{C=1}^C \rho_C \sigma_m^2(C)}{\sigma_m^2} \quad (11)$$

$$\rho_C = \frac{N_C}{\sum_{C=1} N_C} \quad (12)$$

Step 5. Calculate the variance factor v_b^m of d_b^m in the m th feature and the variance factor v_w^m of d_w^m in the m th feature.

$$v_b^m = \frac{\max\|\bar{g}_i^m - \bar{g}_C^m\|}{\min\|\bar{g}_i^m - \bar{g}_C^m\|} \quad (13)$$

$$\|\bar{g}_i^m - \bar{g}_C^m\| = \frac{\bar{g}_i^m - \bar{g}_C^m}{\sigma_m^2} \quad (14)$$

$$v_w^m = \frac{\max(d_w^m)}{\min(d_w^m)} \quad (15)$$

Step 6. Calculate the compensation factor of the m th feature.

$$\eta^m = \frac{1}{v_w^m} + \frac{1}{v_b^m} \quad (16)$$

Step 7. Calculate the distance discrimination factor of the m th feature.

$$\lambda^m = d_b^m - \eta_m d_w^m \quad (17)$$

Step 8. Normalize the distance discriminant factor.

$$\lambda^{m'} = \frac{\lambda^m - \min(\lambda^m)}{\max(\lambda^m) - \min(\lambda^m)} \quad (18)$$

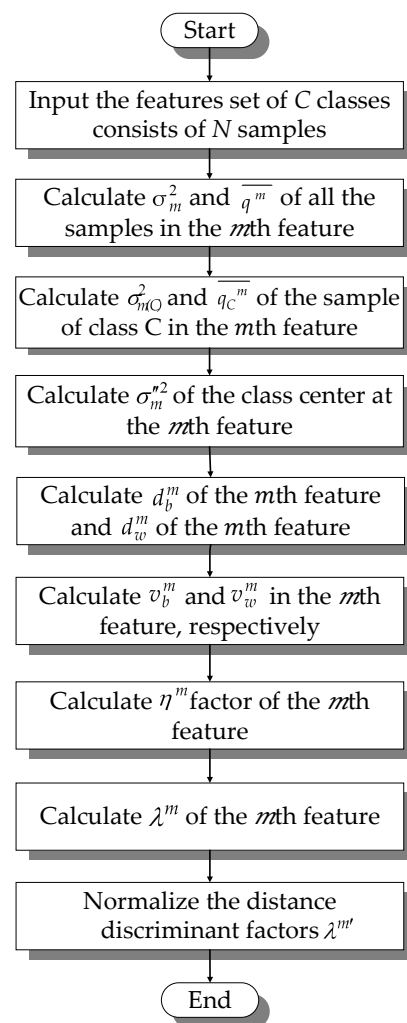


Figure 3. Flow chart of the FSDD.

3.3. Differential Evolution

The differential evolution proposed by Storn and Price is an optimization technology used to solve various complex problems [31]. The calculation principle is similar to the genetic algorithm (GA), including three mechanisms of mutation, crossover and selection. The offspring are derived from random parental parameter mutations, as shown in Figure 4. At the same time, this algorithm refers to particle swarm optimization (PSO) to make the evolution direction approach the best particle. DE is a random search algorithm, and the randomness used in this algorithm prevents the algorithm from falling into the local optimum. Therefore, it can be used for many important problems that need to be optimized, including neural network training, and Bayesian network inference [44]. The algorithms are combined with the DE algorithm to improve the computational efficiency or improve the recognition rate [45,46]. In this article, the identification result was set as the fitness value, the input was the feature rank and the output was the new feature ranking, and the feature ranking order was optimized by DE after features were ranked to improve the identification. The calculation procedure of the differential evolution algorithm is as follows:

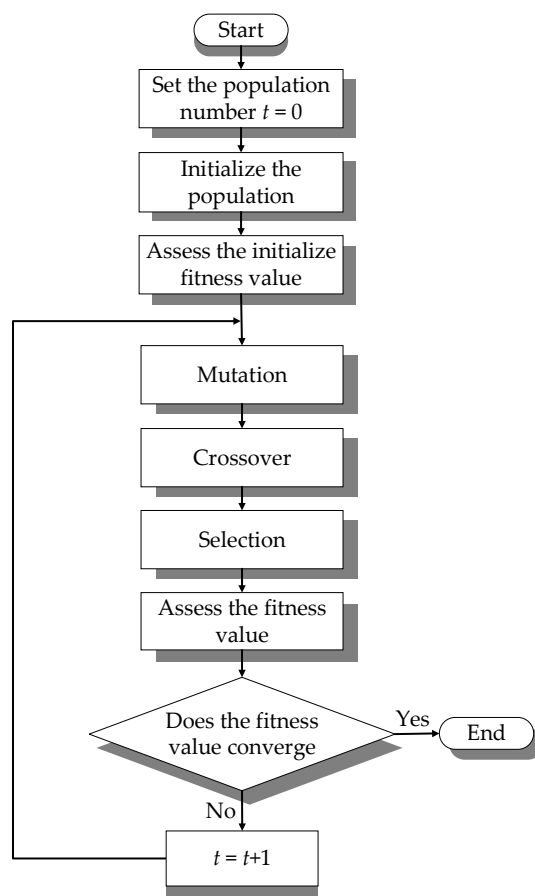


Figure 4. Flow chart of the DE.

Step 1. Initially, set the parameters including the number of particles and the number of iterations j . $G_{1,0}$ is the first generation input of the first particle of the particle group initially.

Step 2. Calculate the fitness value of the first generation of the first particle.

Step 3. Randomly select the parameters in the offspring $G_{1,j}$, $G_{2,j}$ and $G_{3,j}$ to produce mutations.

$$V_{r,j+1} = G_{1,j} + F(G_{2,j} - G_{3,j}) \quad (19)$$

Step 4. The step of crossover is a random operation that bases on *rand* and *CR*. If *CR* is smaller, it means vectors *U* and *G* are more similar.

$$U_{r,j+1} = \begin{cases} V_{r,j+1} & \text{if } rand \leq CR \\ G_{r,j+1} & \text{if } rand > CR \end{cases} \quad (20)$$

Step 5. The step of elimination obtains a better fitness value through the greedy algorithm.

$$U_{r,j+1} = \begin{cases} U_{r,j+1} & \text{if } U_{r,j+1} \leq G_{r,j} \\ G_{r,j} & \end{cases} \quad (21)$$

Step 6. The stopping rule is whether the fitness value has converged, meaning the optimal value. The fitness value is one minus the accuracy rate (1-ACC). If the number of calculations reaches the iterations j , it stops. Otherwise, repeat steps 3 to 5.

Step 7. Finally, all particles converge to obtain the best global solution. After the optimization, a set of solutions can be obtained as the best particle coordinate G_{best} which is the optimized importance of the feature.

3.4. Classifier

The feature rank is obtained by feature selection, which is conducive to the classifier of the neural network model. The nonlinear classifier BPNN will be compared with the linear projection classifier LDA. The classifier randomly selects 70% of the data from the motor samples, and the features are brought into the classifier for training, and the remaining 30% of data are used as test samples.

3.4.1. Backpropagation Neural Network

A BPNN imitates the capabilities of neural system resource processing data and discriminant analysis by simulating the structure of biological data processing. Among them, neurons are used for message transmission and backpropagation to correct errors in order to achieve the best identification result. A BPNN has three structures that are composed of an input layer, a hidden layer and an output layer. The model structure is shown in Figure 5. By entering the input value $X = (X_1, X_2, \dots, X_n)$ and weights $W = (W_1, W_2, \dots, W_n)$ in the input layer, the expected output result $Y = (Y_1, Y_2, Y_3, Y_4)$ will be obtained after the hidden layer $H = (H_1, H_2, H_3, \dots, H_k)$. Then, the error will be corrected by backpropagation, and the new weight will be brought into the forward propagation to obtain the output result $O = (O_1, O_2, O_3, O_4)$.

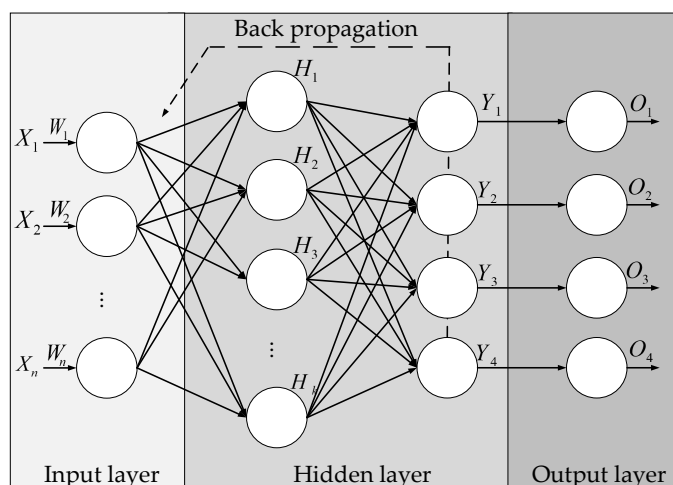


Figure 5. Schematic diagram of a BPNN.

3.4.2. Linear Discriminant Analysis

LDA is used to obtain the new coordinate position of the best distance between the classes and minimize the distance within the class through the projection coordinate axis, which is used to judge different classes. First, one must calculate the intra-class distance matrix S_w and the inter-category distance matrix S_b as

$$S_w = \frac{1}{N} \sum_{c=1}^C \sum_{x \in X_N} (x - \mu_c)(x - \mu_c)^T \tag{22}$$

$$S_b = \frac{1}{C} \sum_{c=1}^C (\mu_c - \mu)(\mu_c - \mu)^T \tag{23}$$

Secondly, the projection matrix $W = S_w^{-1} S_b$ must be calculated by S_w and S_b , where the maximum W is the final projection matrix, and the sample features and projection matrix should be converted into new sample features for classification.

4. Method Efficiency and Robustness

4.1. Dataset Results

We unified the UCI dataset identification results when the initial factor was the same [47]. Then, the feature rank of the dataset was optimized so the features were sequentially entered into the classifier. From Table 3, it can be observed that the identification rate of the BPNN can be significantly improved when the factors are optimized by DE.

Table 3. The accuracy rate of the different recognition systems.

Dataset	Optimize	Classifier	Number of Feature	Accuracy (%)
segmentation	–	BPNN	19	85.62
	DE	BPNN		91.77
	–	LDA		78.35
	DE	LDA		78.42
sonar	–	BPNN	60	84.66
	DE	BPNN		87.21
	–	LDA		71.20
	DE	LDA		72.24
wine	–	BPNN	13	77.96
	DE	BPNN		89.81
	–	LDA		98.30
	DE	LDA		98.90
vowel	–	BPNN	10	44.52
	DE	BPNN		49.61
	–	LDA		58.10
	DE	LDA		60.60
WDBC	–	BPNN	30	62.74
	DE	BPNN		66.98
	–	LDA		95.30
	DE	LDA		95.63

4.2. Identification System Validation Results

The work in this article built a system, and an original signal was brought into the system. In the classifier part of the system, the signal was separately brought into the BPNN and LDA to compare different classifiers which can obtain better results in BLDC fault diagnosis. The features from 2100 samples of data were brought into the classifier for training, whilst the features from the remaining 900 samples of data were used as test samples, and this was repeated 100 times to calculate the average accuracy rate in order to know the resolution of the degree of the type of motor failure. Initially, the input was 96 features, and then the number of inputs decreased after the feature selection. The data matrix was 96×3000 , which means 96 features and 3000 samples from four types of motor.

The signal was directly recognized by the classifier after the feature analysis by the HHT. Although the number of features was the largest, there may be more features that cannot clearly distinguish the fault, which led to an accuracy rate for the BPNN of 95.70%, and an accuracy rate for LDA of 74.89%. Feature selection can reduce computational costs by determining the features of less influence and redundancy.

In Figure 6, the accuracy rate is 93.96% when the FSDD selects 14 features. When the data are brought into the linear classifier LDA, the best accuracy rate is 74.57% in the FSDD feature selection, as shown in Figure 6. Of the two methods, the BPNN can obtain a higher recognition rate and smooth accuracy rate, as when the number of features is 14 in the BPNN, the accuracy rate is 93.96%, as shown in Table 4. From Table 3, it can be observed that the identification rate of the BPNN can be significantly improved when the factors are optimized by DE. However, the result of LDA would not obviously improve. Therefore, it is concluded that the BPNN is a better classifier in this identification system.

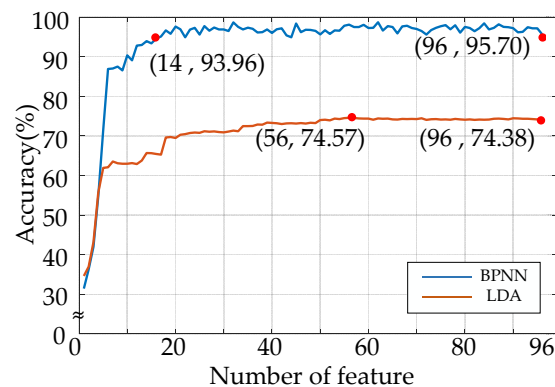


Figure 6. Accuracy curves of the BPNN and LDA.

Table 4. The accuracy rate of the better recognition results in two systems.

Signal Analysis	Feature Selection	Classifier	Number of Features	Accuracy (%)
HHT	FSDD	BPNN	14	93.96
HHT	FSDD	LDA	56	74.57

In Figure 6, we know that the accuracy rate starts to stabilize when the number of features is 14. From this finding, the work in this article used a differential evolution to optimize the factor of the first 14 important features. The first 14 features are F36, F33, F81, F18, F61, F57, F87, F94, F62, F93, F59, F12, F37 and F28.

This research used DE to optimize the factor of the first 14 features that were introduced into the BPNN and LDA. The accuracy rate was increased from 93.96% to 96.00%, and the recognition rate was improved by 3%, as shown in Table 5. The factor of features selected by the FSDD that were optimized by DE can effectively increase the recognition rate and reduce the number of features by 85%.

Table 5. The accuracy rate of the different recognition systems.

Signal Analysis	Feature Selection	Optimizer	Classifier	Number of Features	Accuracy (%)
HHT	–	–	BPNN	96	95.70
HHT	FSDD	–	BPNN	14	93.96
HHT	FSDD	DE	BPNN	14	96.00

The systems conducted signal analysis by the DWT, and the features that were brought into the various classifiers [48] compare with those of the identification system presented in this article. The result of the initial system that just has the HHT and the BPNN was 95.70%, already better than the other systems. The result of the systems established from a variety of classifiers with the DWT were not higher than 90%, and the result of the identification system proposed in this article was 96.00%, as shown in Table 6.

Table 6. The accuracy rate of the different recognition systems.

Signal Analysis	Feature Selection	Optimizer	Classifier	Number of Features	Accuracy (%)
HHT	–	–	BPNN	96	95.70
HHT	FSDD	DE	BPNN	14	96.00
DWT	–	–	Fine Gaussian SVM [48]	112	74.40
DWT	–	–	Fine KNN [48]	112	76.30
DWT	–	–	Bagged Trees [48]	112	89.90
DWT	–	–	Subspace KNN [48]	112	76.30

The Hall signal was analyzed by the HHT, and then the FSDD was used for feature selection, which was combined with DE to optimize the importance of features. Finally, features were brought into the BPNN to obtain the identification result. In Table 7, the original signal is added with different proportions of white noise. After the feature importance was optimized, the noise-free accuracy rate was 96.00%, which is better than the non-optimized accuracy rate of 93.96%. The accuracy rate was 92.04% when the SNR was 20 dB, which is still better than the non-optimized accuracy rate of 90.84% when the SNR was 20 dB. It is confirmed that this method has a robust anti-noise ability.

Table 7. The accuracy rate in the different SNRs.

Signal Analysis	Feature Selection	Optimizer	Classifier	Number of Features	∞ dB	30 dB	25 dB	20 dB
HHT	–	–	BPNN	14	93.96	92.66	92.13	90.84
HHT	FSDD	DE	BPNN		96.00	94.28	92.42	92.04

5. Conclusions

In fault types, bearing damage, stator winding failure and rotor damage make up the majority. The complexity of fault detection is reduced through this system in executing the preliminary diagnosis. This article presented a fault classification system for BLDCs. This system includes four subsystems which are signal analysis, feature selection, ranking optimization and classifiers. In this article, the proposed system reduced the number of features to 14, which significantly eliminated 85% of the redundant features, and the final accuracy rate reached 96.00%, which is higher than the result of the identification of 96 features. In terms of the anti-noise ability, the original signal was added with different white Gaussian noise SNRs = 20 dB. The accuracy rate was 92.04% in the same 14 features; thus, the identification system has a robust anti-noise ability.

Author Contributions: Conceptualization, C.-Y.L. and C.-H.H.; methodology, C.-Y.L. and C.-H.H.; software, C.-Y.L. and C.-H.H.; validation, C.-Y.L. and C.-H.H.; formal analysis, C.-Y.L. and C.-H.H.; investigation, C.-Y.L. and C.-H.H.; resources, C.-Y.L. and C.-H.H.; data curation, C.-Y.L. and C.-H.H.; writing-original draft preparation, C.-Y.L. and C.-H.H.; writing-review and editing, C.-Y.L. and C.-H.H.; visualization, C.-Y.L. and C.-H.H.; 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

- Choi, J.H.; Chun, Y.D.; Han, P.W.; Kim, M.J.; Koo, D.H.; Lee, J.; Chun, J.S. Design of high power permanent magnet motor with segment rectangular copper wire and closed slot opening on electric vehicles. *IEEE Trans. Magn.* **2010**, *46*, 2070–2073. [[CrossRef](#)]
- Usman, A.; Rajpurohit, B.S. Time-efficient fault diagnosis of a BLDC motor drive deployed in electric vehicle applications. In Proceedings of the 2020 IEEE Global Humanitarian Technology Conference (GHTC), Seattle, WA, USA, 29 October–1 November 2020; pp. 2377–6919.
- Tran, M.K.; Bhatti, A.; Vrolyk, R.; Wong, D.; Panchal, S.; Fowler, M.; Fraser, R. A Review of Range Extenders in Battery Electric Vehicles: Current Progress and Future Perspectives. *World Electr. Veh. J.* **2021**, *12*, 54. [[CrossRef](#)]
- Panchal, S. Experimental Investigation and Modeling of Lithium-Ion Battery Cells and Packs for Electric Vehicles. Ph.D. Thesis, University of Ontario Institute of Technology, Oshawa, ON, Canada, 2016.
- Panchal, S. Impact of Vehicle Charge and Discharge Cycles on the thermal Characteristics of Lithium-Ion Batteries. Master's Thesis, University of Waterloo, Waterloo, ON, Canada, 2014.
- Akhoundzadeh, M.H.; Panchal, S.; Samadani, E.; Raahemifar, K.; Fowler, M.; Fraser, R. Investigation and simulation of electric train utilizing hydrogen fuel cell and lithium-ion battery. *Sustain. Energy Technol. Assess.* **2021**, *46*, 101234.

7. Viaene, J.D.; Verbelen, F.; Derammelaere, S.; Stockman, K. Energy-efficient sensorless load angle control of a BLDC motor using sinusoidal currents. *IET Electr. Power Appl.* **2018**, *12*, 1378–1389. [[CrossRef](#)]
8. Mousmi, A.; Abbou, A.; Houm, Y.E. Binary diagnosis of hall effect sensors in brushless dc motor drives. *IEEE Trans. Power Electron.* **2020**, *35*, 3859–3868. [[CrossRef](#)]
9. Zhang, Q.; Feng, M. Fast fault diagnosis method for hall sensors in brushless DC motor drives. *IEEE Trans. Power Electron.* **2019**, *34*, 2585–2596. [[CrossRef](#)]
10. Grubic, S.; Aller, J.M.; Lu, B.; Habetler, T.G. A survey on testing and monitoring methods for stator insulation systems of low-voltage induction machines focusing on turn insulation problems. *IEEE Trans. Ind. Electron.* **2008**, *55*, 4127–4136. [[CrossRef](#)]
11. Shamsi-Nejad, M.A.; Nahid-Mobarakeh, B.; Pierfederici, S.; Meibody-Tabar, F. Fault tolerant and minimum loss control of double-star synchronous machines under open phase conditions. *IEEE Trans. Ind. Electron.* **2008**, *55*, 1956–1965. [[CrossRef](#)]
12. Zidani, F.; Diallo, D.; Benbouzid, M.E.H.; Nait-Said, R. A fuzzy-based approach for the diagnosis of fault modes in a voltage-fed PWM inverter induction motor drive. *IEEE Trans. Ind. Electron.* **2008**, *55*, 586–593. [[CrossRef](#)]
13. Rajagopalan, S.; Aller, J.M.; Restrepo, J.A.; Habetler, T.G.; Harley, R.G. A analytic-wavelet-ridge-based detection of dynamic eccentricity in brushless direct current (BLDC) motors functioning under dynamic operating conditions. *IEEE Trans. Ind. Electron.* **2007**, *54*, 1410–1419. [[CrossRef](#)]
14. Roux, W.I.; Harley, R.G.; Habetler, T.G. Detecting rotor faults in low power permanent magnet synchronous machines. *IEEE Trans. Power Electron.* **2007**, *22*, 322–328. [[CrossRef](#)]
15. Kang, M.; Kim, J.; Kim, J.M. High-performance and energy-efficient fault diagnosis using effective envelope analysis and denoising on a general-purpose graphics processing unit. *IEEE Trans. Power Electron.* **2015**, *30*, 2763–2776. [[CrossRef](#)]
16. Kang, M.; Kim, J.; Kim, J.M.; Tan, A.C.C.; Kim, E.Y.; Choi, B.K. Reliable fault diagnosis for low-speed bearings using individually trained support vector machines with kernel discriminative feature analysis. *IEEE Trans. Power Electron.* **2015**, *30*, 2786–2797. [[CrossRef](#)]
17. Ebrahimi, B.M.; Faiz, J.; Roshtkhari, M.J. Static-, dynamic-, and mixed-eccentricity fault diagnoses in permanent-magnet synchronous motors. *IEEE Trans. Ind. Electron.* **2009**, *56*, 4727–4739. [[CrossRef](#)]
18. Zhang, J.H.; Zhao, J.; Zhou, D.; Huang, C. High-performance fault diagnosis in PWM voltage-source inverters for vector-controlled induction motor drives. *IEEE Trans. Power Electron.* **2014**, *11*, 6087–6099. [[CrossRef](#)]
19. Nandi, S.; Toliyat, H.A.; Li, X. Condition monitoring and fault diagnosis of electrical motors—A review. *IEEE Trans. Energy Convers.* **2005**, *20*, 719–729. [[CrossRef](#)]
20. Herrera, A.L.M.; Carrillo, L.M.L.; Ramirez, M.L.; Colores, S.S.; Yepez, E.C. Gabor and the Wigner-Ville transforms for broken rotor bars detection in induction motors. In Proceedings of the International Conference on Electronics, Communications and Computers, Cholula, Mexico, 26–28 February 2014.
21. Osman, S.; Wang, W. A morphological hilbert-huang transform technique for bearing fault detection. *IEEE Instrum. Meas. Mag.* **2016**, *65*, 2646–2656. [[CrossRef](#)]
22. Goharrizi, A.Y.; Sepehri, N. Internal leakage detection in hydraulic actuators using empirical mode decomposition and hilbert spectrum. *IEEE Trans. Instrum. Meas.* **2012**, *61*, 368–378. [[CrossRef](#)]
23. Song, Q.; Ni, J.; Wang, G. A fast clustering-based feature subset selection algorithm for high-dimensional data. *IEEE Trans. Knowl. Data Eng.* **2013**, *25*, 1–14. [[CrossRef](#)]
24. Estevez, P.A.; Tesmer, M.; Perez, C.A.; Zurada, J.M. Normalized mutual information feature selection. *IEEE Trans. Neural Netw.* **2009**, *20*, 189–201. [[CrossRef](#)]
25. Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 1226–1238. [[CrossRef](#)]
26. Lin, F.J.; Chen, C.I.; Lin, J.R. Detection of mechanical resonance frequencies for interior permanent magnet synchronous motor servo drives based on wavelet multiresolution filter. *IET J. Eng.* **2020**, *2020*, 827–833. [[CrossRef](#)]
27. Liu, X.Y.; Liang, Y.; Wang, S.; Yang, Z.Y.; Ye, H.S. A hybrid genetic algorithm with wrapper-embedded approaches for feature selection. *IEEE Access* **2018**, *6*, 22863–22874. [[CrossRef](#)]
28. Liu, H.; Yu, L. Toward integrating feature selection algorithms for classification and clustering. *IEEE Trans. Knowl. Data Eng.* **2005**, *17*, 491–502.
29. Liu, C.; Wu, C.; Jiang, L. Evolutionary clustering framework based on distance matrix for arbitrary-shaped data sets. *IET Signal Process.* **2016**, *10*, 478–485. [[CrossRef](#)]
30. Liang, J.; Yang, S.; Winstanley, A. Invariant optimal feature selection: A distance discriminant and feature ranking based solution. *Pattern Recognit.* **2008**, *41*, 1429–1439. [[CrossRef](#)]
31. Fan, Q.; Yan, X. Self-adaptive differential evolution algorithm with zoning evolution of control parameters and adaptive mutation strategies. *IEEE Trans. Cybern.* **2016**, *46*, 219–232. [[CrossRef](#)]
32. Filippetti, F.; Franceschini, G.; Tassoni, C.; Vas, P. Recent developments of induction motor drives fault diagnosis using AI techniques. *IEEE Trans. Ind. Electron.* **2000**, *47*, 994–1004. [[CrossRef](#)]
33. Chen, D.; Liu, Y.; Zhou, J. Optimized neural network by genetic algorithm and its application in fault diagnosis of three-level inverter. In Proceedings of the 2019 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS), Xiamen, China, 5–7 July 2019.

34. Lin, J.W.; Chao, C.T.; Chiou, J.S. Determining neuronal number in each hidden layer using earthquake catalogues as training data in training an embedded back propagation neural network for predicting earthquake magnitude. *IEEE Access* **2018**, *6*, 52582–52597. [[CrossRef](#)]
35. Amrutha, J.; Ajai, A.S.R. Performance analysis of backpropagation algorithm of artificial neural networks in verilog. In Proceedings of the 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 18–19 May 2018.
36. Abohagar, A.A.; Mustafa, M.W. Back propagation neural network aided wavelet transform for high impedance fault detection and faulty phase selection. In Proceedings of the 2012 IEEE International Conference on Power and Energy (PECon), Kota Kinabalu, Malaysia, 2–5 December 2012.
37. Gaxiola, F.; Melin, P.; Valdez, F.; Castillo, O. Backpropagation learning method with interval type-2 fuzzy weights in neural networks. In Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX, USA, 4–9 August 2013.
38. Montesdeoca-Contreras, J.C.; Zambrano-Abad, J.C.; Morales-Garcia, J.A.; Ávila-Campoverde, R.S. Virtual speed sensor for dc motor using back-propagation artificial neural networks. In Proceedings of the 2014 IEEE international autumn meeting on power, Electronics and Computing (ROPEC), Ixtapa, Mexico, 5–7 November 2014.
39. Niu, G.; Son, J.D.; Widodo, A.; Widodo, A.; Yang, B.S.; Hwang, D.H.; Kang, D.S. A comparison of classifier performance for fault diagnosis of induction motor using multi-type signals. *Struct. Health Monit.* **2007**, *6*, 215–229.
40. Dhir, C.S.; Lee, S.Y. Discriminant independent component analysis. *IEEE Trans. Neural Netw.* **2011**, *22*, 845–857. [[CrossRef](#)] [[PubMed](#)]
41. Stankovic, L.; Thayaparan, T.; Dakovic, M. Signal decomposition by using the S-method with application to the analysis of hf radar signals in sea-clutter. *IEEE Trans. Signal Process.* **2006**, *54*, 4332–4342. [[CrossRef](#)]
42. Kowalski, M.; Torresani, B. Random models for sparse signals expansion on unions of bases with application to audio signals. *IEEE Trans. Signal Process.* **2008**, *56*, 3468–3481. [[CrossRef](#)]
43. Zão, L.; Coelho, R. On the estimation of fundamental frequency from nonstationary noisy speech signals based on the Hilbert–Huang transform. *IEEE Signal Process Lett.* **2018**, *25*, 248–252. [[CrossRef](#)]
44. Strasser, S.; Sheppard, J.; Fortier, N.; Goodman, R. Factored Evolutionary Algorithms. *IEEE Trans. Evol. Comput.* **2017**, *21*, 281–293. [[CrossRef](#)]
45. Hu, K.; Liu, Z.; Huang, K.; Dai, C.; Gao, S. Improved differential evolution algorithm of model-based diagnosis in traction substation fault diagnosis of high-speed railway. *IET Electr. Syst. Transp.* **2016**, *6*, 163–169. [[CrossRef](#)]
46. Secmen, M.; Tasgetiren, M.F. Ensemble of differential evolution algorithms for electromagnetic target recognition problem. *IET Radar Sonar Navig.* **2013**, *7*, 780–788. [[CrossRef](#)]
47. UCI Machine Learning Repository. Available online: <http://archive.ics.uci.edu/ml> (accessed on 5 September 2019).
48. Ali, M.Z.; Shabbir, M.N.S.K.; Liang, X.; Zhang, Y.; Hu, T. Machine learning-based fault diagnosis for single- and multi-faults in induction motors using measured stator currents and vibration signals. *IEEE Trans. Ind. Appl.* **2019**, *55*, 2378–2391. [[CrossRef](#)]