

RESEARCH ARTICLE

Intelligent Fault Diagnosis for BLDC With Incorporating Accuracy and False Negative Rate in Feature Selection Optimization

CHUN-YAO LEE¹, (Member, IEEE), CHEN-HSU HUNG¹, AND TRUONG-AN LE²¹Department of Electrical Engineering, Chung Yuan Christian University, Taoyuan 320314, Taiwan²Department of Electrical and Electronic Engineering, Thu Dau Mot University, Thu Dau Mot, Binh Duong 75000, Vietnam

Corresponding author: Chun-Yao Lee (cyl@cycu.edu.tw)

ABSTRACT Early fault diagnosis is essential for the proper operation of rotating machines. This article proposes a fitness function in differential evolution (DE) that considers accuracy rate and false negative rate for optimization in brushless DC (BLDC) motor fault diagnosis. Feature selection based on a distance discriminant (FSDD) calculates the feature factors which base on the category separability of features after the Hilbert–Huang transform (HHT) which extracts the features of four different type signals from BLDC motor Hall sensor. The feature rank through DE to optimize before the features into the backpropagation neural network (BPNN) in order. By reducing the feature number of Hall signal and decreasing the complexity of neural network input, the combined method was proposed in this article can significantly reduce the calculation cost. Finally, the identification model obtained an accuracy rate of 98.98% and false negative rate of 13.66% when there were 18 features; besides, receiver operating characteristic (ROC) curve and probability curve have been evidenced the number of false negative is decreased. Moreover, the experiments have verified that the proposed method is effective in UCI data set.


INDEX TERMS Brushless DC, differential evolution, fault diagnosis, feature selection.

I. INTRODUCTION

Intelligent fault diagnosis (IFD), applications of machine learning theories to machine fault diagnosis, is a promising way to release the contribution from human labor and automatically recognize the health states of machines. A commonly-implemented diagnosis procedure includes feature extraction, feature selection and classifiers. Deep learning (DL) is compared with IFD which features are manually extracted and selected with required prior knowledge. With the purpose of avoiding manual feature extraction, DL has been successfully applied to fault diagnosis, including a fault diagnosis model is designed with a novel stacked transfer auto-encoder (NSTAE) [1], an adaptive feature learning approach that is built on spatiotemporal pattern network (STPN) [2], and deep belief network (DBN) be used as the classification [3]. DL through adjusting connected weights

to learn and select representations and patterns that can best represent the condition of working machines from the input data [4]. It helps automatically learn fault features from the collected data instead of the artificial feature extraction of IFD. However, the computational cost for deep learning models is high because the input data is usually high dimensional network structure. And these models require a great deal of training data to be labeled, even only a fraction of condition monitoring data is labeled in the real issues [5].

The work in this article constructed a model 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 [6]. A Hall sensor has the obvious advantages of a low cost and simple structure [7]. Additionally, DC motors using Hall sensors have been widely used in commercial and industrial applications [8]. Therefore, this article uses the Hall signal, which is an electrical technology, to establish an identification model. The motor may suffer from different failures, including stator

The associate editor coordinating the review of this manuscript and approving it for publication was Jon Atli Benediktsson .

failure [9]–[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].

The comprehensive description of the proposed fault diagnosis model includes feature extraction, feature selection 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]. Since there does not have to choose the mother wavelet, it is not affected by the resolution of the time domain and frequency domain, so this method can more accurately decompose the signal in the high frequency domain.

After signal analysis, the original signal can provide features with a good identification rate through feature extraction and feature selection. Feature selection can be divided into filter, wrapper, hybrid approach and embedded feature selection [21]. The filter type is based on the relationship between features as the criterion [22], and the wrapper type is based on the relationship between features and the target variable as the criterion [23]. The embedded type is usually used for high dimensional data features [24], [25]. The wrapper and hybrid approach types can obtain better results [26], but the filter type is usually used when considering the computational cost and a large number of features [24], [25]. Therefore, this study used the filter feature selection to calculate feature weights, such as the Feature selection based on a distance discriminant (FSDD), belonging to a clustering algorithm [27].

The main disadvantage of the filter feature selection is that this method independently examines the relationship between features, and the lack of a classifier to participate in the feature selection process leads to ignoring the performance of the feature for the identification results. In the proposed model, the features with distance discriminant factor from FSDD are optimized by DE, the classifier is involved in the feature selection process through optimization to obtain the features with optimized feature ranking. This article combines DE with feature factors after feature selection to optimize the feature ranking. Differential evolution (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 [28].

An artificial neural network (ANN) is a common nonlinear function processor that imitates the structure and pattern of the human brain [29]. The performance of the learning process of the neural network depends on the weights of the neural network in the training phase. A BPNN is a supervised machine learning technique that adjusts its weights to minimize the error of the calculated output, and it is suitable for identifying nonlinear relationships [30]. A BPNN is used in the fault diagnosis problem of NPC inverters [31],

high impedance faults [32] and virtual speed sensors for DC motors [33].

Performance evaluation by metrics which includes true positive (TP), false positive (FP), true negative (TN) and false negative (FN) in novel convolutional neural network [34]. Regulated parameter to reduce variations that can lead to false alarms in healthy operating conditions of the motor in wound rotor induction machine drives [35]. Error probability (false alarm and miss alarm), one of the thresholds, determine would continue iterative decision making in the fault diagnosis scheme [36]. In most of the literature, false alarm or miss alarm probability were mentioned to exactly identify the performance of the proposed fault diagnosis scheme.

Based on the abovementioned related literature, this research proposes a fault identification model for a BLDC established by Hall signals, which includes signal analysis selection, feature selection and classifiers. Besides, that discusses the number of false positive in the model to find out the performance shift of the proposed fault diagnosis scheme.

II. EXPERIMENTAL SETUP AND RELATED METHOD

A. EXPERIMENTAL ARCHITECTURE

This section introduces the experimental equipment, experimental architecture and signal samples in this research and studies the healthy, bearing, winding and rotor, a total of four different types in a BLDC to build a fault diagnosis model. The process of this research is that the servo motor (11kW/2000rpm/69Hz) of the dynamometer generates the opposite torque to the BLDC (420W/3020 rpm/DC 24V/60Hz) as the load, and then the BLDC motor drives the operation. The BLDCM parameters are listed in Table 1. A total of four BLDCs were tested in this experiment. One motor was healthy, 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.

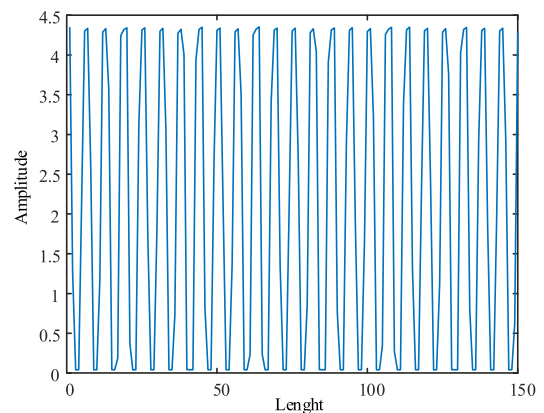


FIGURE 1. Original output voltage of the healthy BLDC hall-sensor in 150 points.

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 %

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 seconds. There was a total of 1500 seconds of measurement records for BLDC motors in each condition, and the 1500 second data were divided into 750 samples of data, every sample having 2000 points. The adopted procedure use HHT be preprocessing tools of the healthy BLDC hall-sensor signals, the original output voltage of the hall-sensor in 150 points is shown in Fig. 1. Matlab program is used to compile and analyze signals in the personal computer with Intel Core i5-4460 3.2 GHz and 8 GB RAM after the Hall signal of the DC brushless motor was acquired by the data acquisition system (NI PXIe-1073).

A total of 3000 samples of Hall signal data records for motors 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. 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 processing and configuration are shown in Fig. 2.

B. SIGNAL ANALYSIS AND FEATURE EXTRACTION

The signal can be analyzed 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 [37]. 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 [38]. HHT consists of empirical mode decomposition (EMD) and Hilbert transform (HT). The original function of the input can be decomposed into intrinsic mode functions (IMFs) and trend functions through EMD. EMD which is a series of shifting decomposition of signal, a highly efficient data decomposition method, adaptively extract the basis functions from signals. Hilbert spectral analysis that each IMF is analyzed by Hilbert transform which transfer to plural modus to get the Hilbert spectrum of the signal. EMD separates the four types of motor Hall signals into the eight layers (IMF1 to IMF7 and Residual), and Hilbert–Huang transform decompose the IMFs to obtain the instantaneous amplitude and instantaneous frequency of each layer. Additionally, there have 12 features that are captured are 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 of every single sample 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. The experiment uses feature extraction method to obtain the feature set before optimizing the feature rank by feature selection and optimizer methods and then uses classifiers to present the results, as shown in Fig. 3.

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
	Residual	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
	Residual	F91	F92	F93	F94	F95	F96

C. FEATURE SELECTION

The model includes feature selection in order to eliminate redundant features which improve prediction accuracy and reduce computational cost. The feature selection was implemented using the FSDD of the clustering algorithm to calculate the category separability of features. The higher value of the factor represents which is the more important feature.

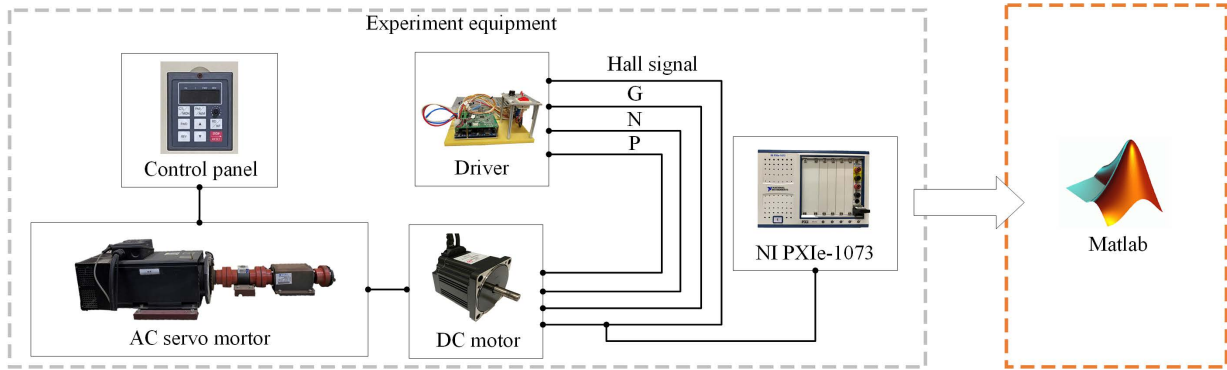


FIGURE 2. Experimental architecture process.

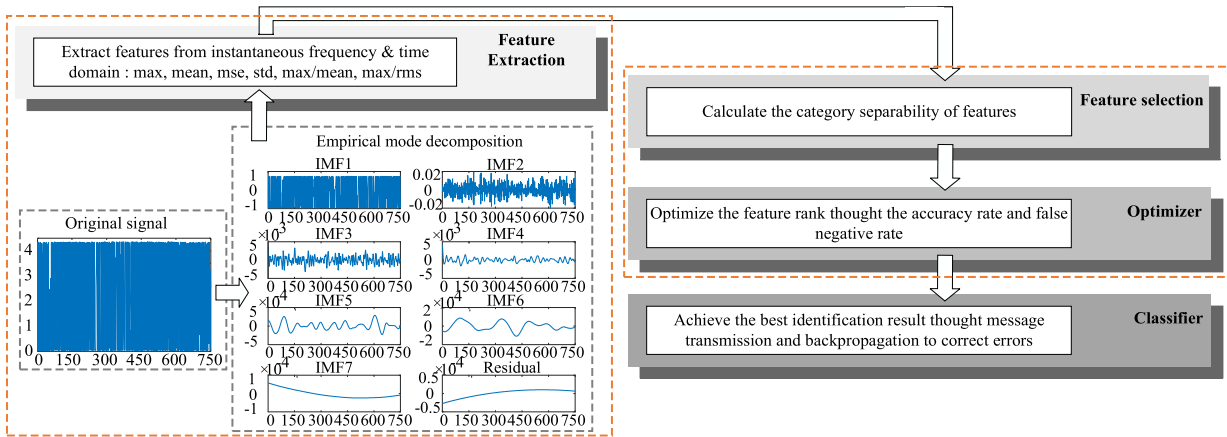


FIGURE 3. Fault diagnosis model architecture.

Therefore, the features were ranked in descending order by feature factors after feature selection. The features can increase the recognition rate of the classifier or not affect the recognition result by feature selection and deletion which are extracted from the Hall signal after signal analysis, which can reduce the computation cost of the recognition model. 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 categories d_b^m . 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.

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

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

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

$$\sigma_{m(C)}^2 = \frac{1}{N_C} \sum_{i=1}^{N_C} (q_i^m - \bar{q}_C^m)^2 \quad (3)$$

$$\bar{q}_C^m = \frac{1}{N_C} \sum_{i=1}^{N_C} q_i^m \quad (4)$$

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

$$\sigma_m'^2 = \sum_{C=1}^C \rho_C (g_C^m - g^m)^2 \quad (5)$$

$$g^m = \sum_{C=1}^C \rho_C g_C^m \quad (6)$$

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

$$d_b^m = \frac{\sigma_m'^2}{\sigma_m^2} \quad (7)$$

$$d_w^m = 2 \frac{\sum_{C=1}^C \rho_C \sigma_{m(C)}^2}{\sigma_m^2} \quad (8)$$

$$\rho_C = \frac{N_C}{\sum_{C=1} N_C} \quad (9)$$

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.

$$v_b^m = \frac{\max \left\| \overline{g_i^m} - \overline{g_C^m} \right\|}{\min \left\| \overline{g_i^m} - \overline{g_C^m} \right\|} \quad (10)$$

$$\left\| \overline{g_i^m} - \overline{g_C^m} \right\| = \frac{\overline{g_i^m} - \overline{g_C^m}}{\sigma_m^2} \quad (11)$$

$$v_w^m = \frac{\max (d_w^m)}{\min (d_w^m)} \quad (12)$$

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

$$\eta^m = \frac{1}{v_w^m} + \frac{1}{v_b^m} \quad (13)$$

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

$$\lambda^m = d_b^m - \eta^m d_w^m \quad (14)$$

Step 8. Normalize the distance discriminant factor.

$$\lambda^{m'} = \frac{\lambda^m - \min (\lambda^m)}{\max (\lambda^m) - \min (\lambda^m)} \quad (15)$$

D. CLASSIFIER

The feature rank is obtained by feature selection, which is conducive to the classifier of the neural network model. The nonlinear classifier BPNN which 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. The size of input layer depends on the feature number, the size of hidden layer is 60 and the training function which is `trainsecg` uses less memory. The BPNN parameters are listed in Table 3. 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.

III. THE PROPOSED METHOD

A. DIFFERENTIAL EVOLUTION

Differential evolution proposed by Price and Storn is an optimization technique that is a competitive and reliable evolutionary used to solve various complex problems [39]. 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 Fig. 4. In addition, this algorithm refers to particle swarm optimization (PSO) to make the evolution direction approach the best particle. The randomness used in DE is a random search algorithm that prevents the

TABLE 3. BPNN Parameters.

Parameters	Value
Hidden Layer Size	60
Output Layer Size	4
Training Function	<code>trainsecg</code>
Training Ratio	70/100
Validation Ratio	15/100
Testing Ratio	15/100
Performance Function	Cross-Entropy

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 [40]. Other articles have proposed the algorithms which were combined with the DE algorithm to improve the computational efficiency or improve the recognition rate [41], [42]. In this article, the accuracy rate and false negative rate are combined to be set as the fitness value, the feature rank is a rank optimized by DE after optimized to improve the identification and false negative rate. The calculation procedure of the differential evolution algorithm is as follows:

Step 1. Initially, set the parameters as follow: the number of population $r = 10$, the numbers of particle which are the numbers of feature $i = 96$, crossover rate $CR=0.8$, $G_{1,0}$, the distance discriminant factor from FSDD, is the first generation of the first particle in population and initialize iterations $j = 0$.

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.

$$V_{r,j+1} = G_{1,j} + F (G_{2,j} - G_{3,j}) \quad (16)$$

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.

$$U_{r,j+1} = \begin{cases} V_{r,j+1} & \text{if } \text{rand} \leq CR \\ G_{r,j+1} & \text{if } \text{rand} > CR \end{cases} \quad (17)$$

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

$$U_{r,j+1} = \begin{cases} U_{r,j+1} & \text{if } U_{r,j+1} \leq G_{r,j} \\ G_{r,j} & \end{cases} \quad (18)$$

Step 6. The stopping rule is whether the fitness value has converged, meaning the optimal value. The fitness value combine the accuracy rate and false negative rate. 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

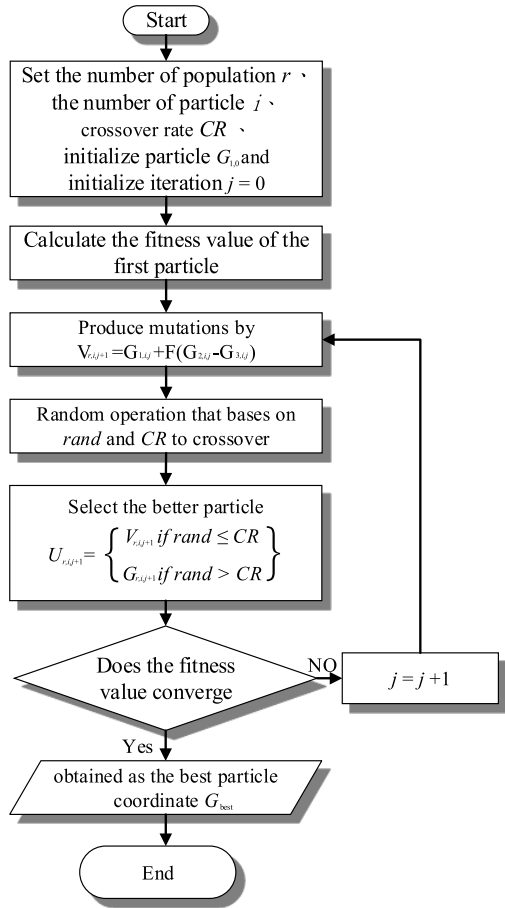


FIGURE 4. Flow chart of the DE.

be obtained as the best particle coordinate G_{best} which is the optimized importance of the feature.

B. FITNESS FUNCTION

There are four possible results in the recognition test, true positive, true negative, false positive and false negative. A false negative (FN) is an error in which a test result is negative of a positive subject, while a false positive (FP) is an opposite error, where the test result is negative subject is positive. The false positive and false negative exist where the negative samples (left curve) and the positive samples (right curve) overlap, as shown in Fig. 5. The false negative rate (FNR) can be decreased by setting different thresholds (vertical bar), at the cost of raising the number of false positive (FP) and decreasing the accuracy rate (ACC), or vice versa [43]. A perfect test would have zero false positives and zero false negatives, so building the curve of samples by the different recognized results is another way. The false negative rate is shown in (20), where FN is the number of false negatives, TP is the number of true positives and $N = FN + TP$ is the total number of ground truth positives [40].

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \times 100 \% \quad (19)$$

$$FNR = \frac{FN}{FN + TP} \times 100 \% \quad (20)$$

In the following paragraph, the equations of fitness value combine the accuracy rate and the false negative rate, and then the property of equations are represented in the surface plot which is a visual medium to exactly describe the performance. The red points in the surface plot of fitness evolution are the accuracy rate and false negative rate before optimization to state the global situation.

Equation 21 referent the Pearson correlation coefficient to combine the accuracy rate and false negative rate after optimizing in the iterations. ACC and FNR are accuracy rate and false negative rate. There has obviously been scaling in $Fitness$ when the FNR is extremely smaller, $Fitness$ is flat at other times. The increased degree of fitness is unobvious that could not search the evolution direction as shown in Fig. 6.

$$Fitness = \frac{k \times ACC}{\sqrt{k + k \times (k - 1) \times FNR}} \quad (21)$$

Equation 22 cancel the division in the equation to increase the degree of change in fitness. \overline{FNR} is original accuracy rate and false negative rate. There have many combined options of ACC and FNR with the same fitness value. It leads to an increase in $Fitness$, while ACC and FNR are not optimized at the same time as shown in Fig. 7.

$$Fitness = ACC + (\overline{FNR} - FNR) \quad (22)$$

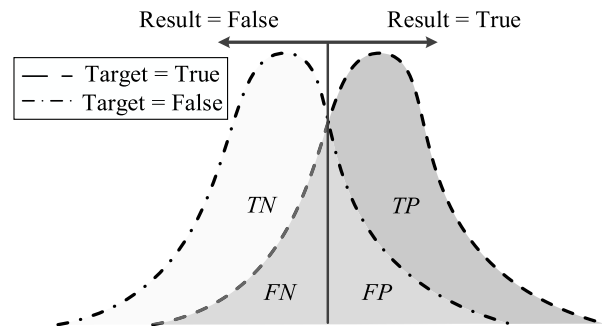


FIGURE 5. Schematic diagram of probability distribution.

Equation 23, $Fitness$ will be improved when the ACC is increased and the FNR is reduced at the same time. \overline{ACC} and \overline{FNR} are original accuracy rate and false negative rate, which are the origin of coordinate in the evolution plot. There has the plain and slope even the ACC and FNR are in the micro shift so that the optimization has a clear direction as shown in Fig. 8.

$$Fitness = |ACC - \overline{ACC}| \times (\overline{FNR} - FNR) + (ACC - \overline{ACC}) \times |\overline{FNR} - FNR| \quad (23)$$

IV. METHOD EFFICIENCY

A. DATASET RESULTS

The work in this article built a model, and an original signal was brought into the model. In the classifier part of the model,

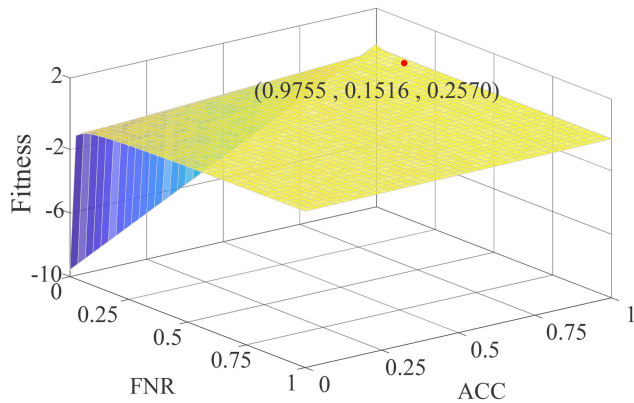


FIGURE 6. Fitness evolution plot of pearson correlation coefficient.

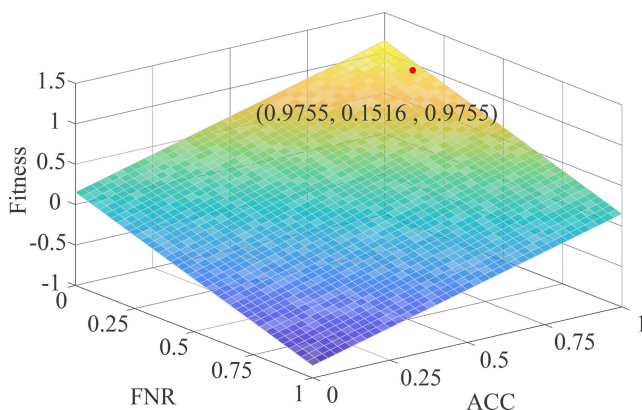


FIGURE 7. Fitness evolution plot of equation 25.

the signal was brought into the BPNN which can obtain better results in BLDC fault diagnosis. The features from 2,100 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%. Feature selection can reduce computational costs by determining the features of less influence and redundancy.

In Fig. 9, the method can obtain a smooth accuracy rate after the number of features was more than 10. The accuracy rate is 95.70% when the number of features is 96; the highest accuracy rate in a full experiment is 99.25% when the number of features is 62; the lowest false negative rate in a full experiment is 1.00% when the number of features is 58.

TABLE 4. The accuracy rate and false negative rate with different number of features.

Number of Features	False Negative Rate	Accuracy Rate
13	15.25	95.74
14	27.82	93.96
15	23.59	94.21
16	16.15	94.57
17	29.72	95.01
18	15.16	97.55
19	36.70	94.09

When number of features is 96, there is no outstanding advantage in accuracy and false negative; the false negative rate is the lowest when there are 58 features but the number of features is large; the recognition rate is the highest when there are 62 features but the false negative rate is high. As mentioned above is the reason that the number of features that will be optimized should be another position. When the recognition rate is around 16 features, it gradually stabilizes. And the false negative rate is also relatively lower in the entire experiment result.

In Fig. 9, the accuracy rate is gradually stabilized from around 10 features, but the false negative rate is unstable even in more features. There has a feature point worth being noteworthy that the false negative rate is relatively lower and the beginning of stabilized accuracy rate when the number of features is 18. From this finding, the work in this article used a differential evolution to optimize the factor of the first 18 important features.

B. OPTIMIZATION

The number of 18 features in Table 4 is the best, and the recognition rate and false negative rate are both the best values. Although the 18 features did not have outstanding advantages in the overall experiment. The 18 features compared with the number of features previously proposed, only need to use a lesser number of features are recognized, and the shortcomings can be optimized to increase the recognition rate and reduce the false negative rate.

In Table 5, when the fitness value combines the accuracy rate and false negative rate as (21) can effectively enhance fitness, greatly reduce false negatives, and reduce the accuracy rate. Adding the original point of fitness into the fitness formula as (22) can effectively improve fitness, greatly reduce false negatives, and slightly reduce the recognition rate. Considering the original point of fitness and mutation direction, (23) effectively increases the recognition rate and reduces false negatives at the same time which is the best fitness function in these three equations.

This subsection uses the continuous dataset provided by UCI for verification [44], experimental verifications are performed to demonstrate the effectiveness of the proposed

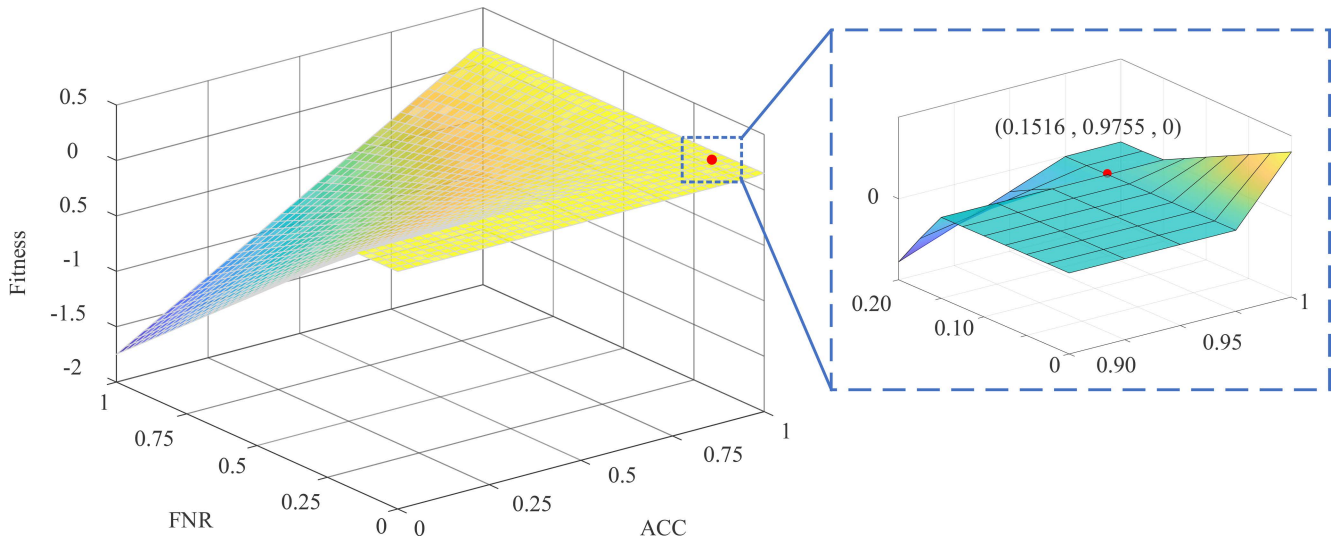


FIGURE 8. Fitness evolution plot of equation 26.

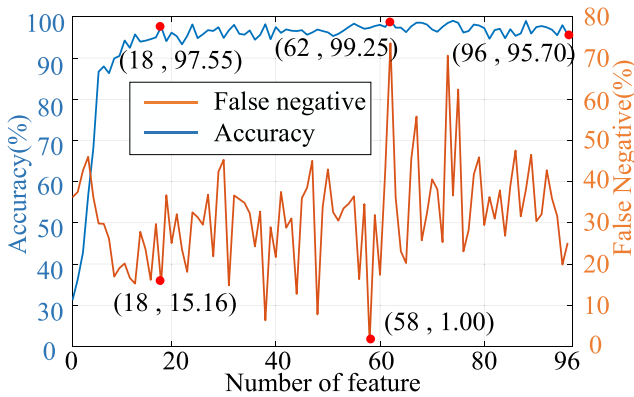


FIGURE 9. Accuracy and false negative curves.

TABLE 5. The accuracy rate and false negative rate with different fitness of optimization.

Fitness	Number of Features	False Negative Rate	Accuracy Rate
—	18	15.16	97.55
(24)	18	7.00	90.33
(25)	18	10.54	95.52
(26)	18	13.66	98.98

method, and the features of the dataset are carried out with the (23) in DE. First, the feature is selected by using the FSDD before ranking the feature factors. Then features are entered into the DE optimizer. The optimizer results are shown in Table 6, the false negative rates and accuracy rates after optimizing were better than the original. ROC curve compares the TPR and FPR in the plot to confirm the FP situation [45], [46]. If only changing result cutoff values could lead the FP to be raised when decreasing the FN [47], [48]. The ROC curves

TABLE 6. The accuracy rate and false negative rate of the different recognition.

Dataset	Number of Features	Optimizer	False Negative Rate	Accuracy Rate
segmentation	19	—	60.54 %	68.74 %
		DE	43.79 %	70.69 %
sonar	60	—	16.66 %	83.34 %
		DE	13.06 %	86.94 %
vowel	13	—	13.74 %	47.06 %
		DE	7.94 %	49.39 %
wine	10	—	5.77 %	93.79 %
		DE	0 %	94.23 %
WDBC	30	—	4.12 %	95.88 %
		DE	1.74 %	98.26 %

for diagnostic with optimization and without optimization are plotted in Fig. 10 that illustrates the FP would not obviously increase in the proposed method. The area under the ROC curve (AUC), a metric derived from the ROC curve, evaluates the performance of classifiers. The AUC values of model with DE are higher than the model without optimization, except for the AUC value of winding short circuit motor which is nuance in the two models, as shown in Table 7. In order to compare the probability curve after optimization, the probability distribution of motor accuracy in four types is created by accuracy and probability. Fig. 11 shows the probability of one type motor being classified as the four conditions of motors. The fault motor probability curve of rotor damage and short circuit in stator windings moves to both ends of accuracy after optimization that expounds the

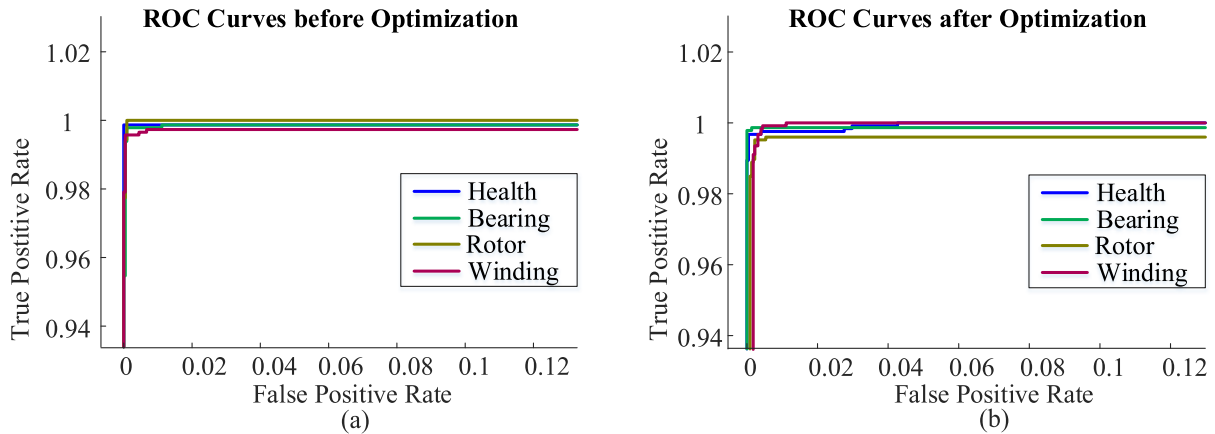


FIGURE 10. ROC curves of four kind motor result (a) before optimization (b) after optimization.

TABLE 7. AUC value of four kind motor.

Optimization	AUC value of health motor	AUC value of bearing damage motor	AUC value of rotor damage motor	AUC value of winding short circuit motor
Without optimization	0.99924	0.99814	0.99856	0.99820
DE optimization	0.99985	0.99955	0.99991	0.99803

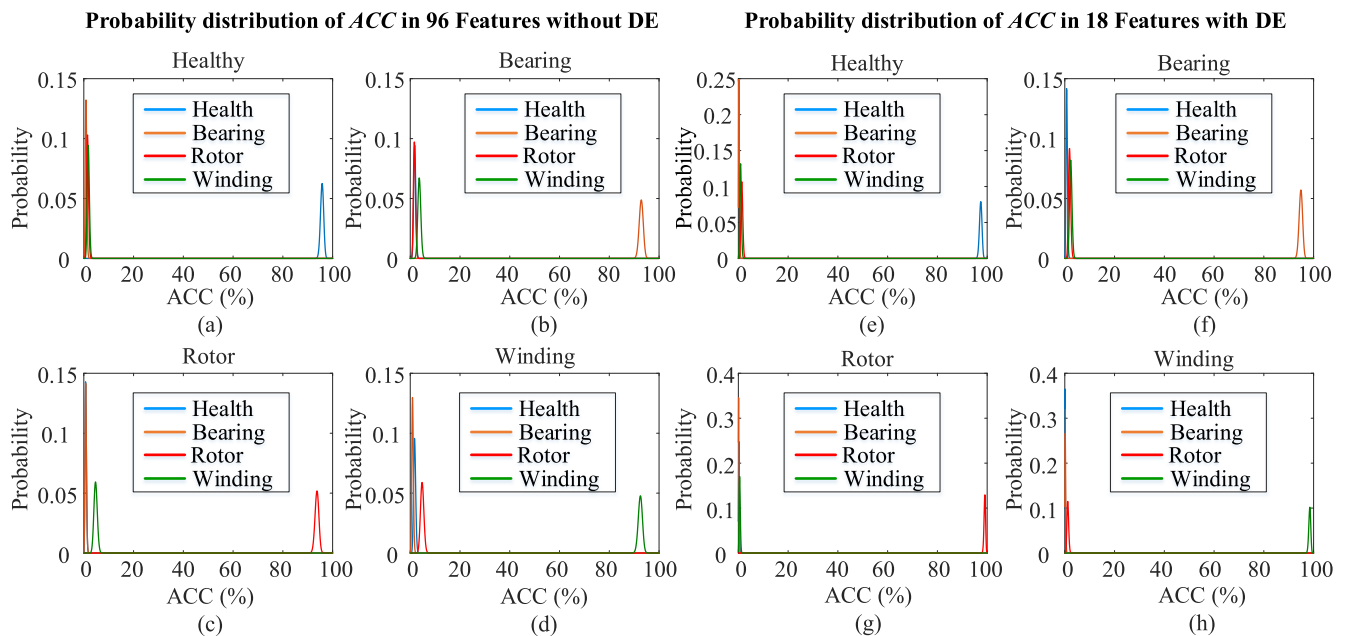


FIGURE 11. Probability distribution of motor accuracy in four types. (a) the identification result of the healthy motor dataset (b) the identification result of the bearing damage motor dataset (c) the identification result of the rotor damage motor dataset (d) the identification result of the winding damage motor dataset in 96 features without DE model. (e) the identification result of the healthy motor dataset (f) the identification result of the bearing damage motor dataset (g) the identification result of the rotor damage motor dataset (h) the identification result of the winding damage motor dataset in 18 features with DE model.

probability of these two types being classified as the correct motor raises, as shown in Fig. 11(c), (d), (g), (h).

ROC curve has argued the false negative numbers of model with the proposed fitness function as similarly less as without optimization. Above ROC result and probability curve has been testified the model with the proposed fitness

function is better than only moving the threshold to raise the accuracy rate.

V. CONCLUSION

In fault types, bearing damage, stator winding failure and rotor damage make up the majority. The complexity

of fault detection is reduced through this model in executing the preliminary diagnosis. This article presented a fault diagnosis model for BLDCs. This model includes five subsystems which are signal analysis, feature extraction, feature selection, ranking optimization and classifiers.

In this article, the proposed model reduced the number of features to 18, which significantly eliminated 81% of the redundant features in the BLDC dataset, and then the model has been successfully applied to the other dataset. The final accuracy rate reached 98.98% and false negative rate was 13.66% in the BLDC dataset, which is higher than the result of the identification of 96 features. Moreover, the model with the proposed fitness function has been testified through the ROC curve and probability curve this is more than moving the threshold and increasing the accuracy probability even.

REFERENCES

- [1] S. Haidong, D. Ziyang, C. Junsheng, and J. Hongkai, "Intelligent fault diagnosis among different rotating machines using novel stacked transfer auto-encoder optimized by PSO," *ISA Trans.*, vol. 105, pp. 308–319, Oct. 2020.
- [2] T. Han, C. Liu, L. Wu, S. Sarkar, and D. Jiang, "An adaptive spatiotemporal feature learning approach for fault diagnosis in complex systems," *Mech. Syst. Signal Process.*, vol. 117, pp. 170–187, Feb. 2019.
- [3] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1693–1702, Jul. 2017.
- [4] S. Shao, R. Yan, Y. Lu, P. Wang, and R. Gao, "DCNN-based multi-signal induction motor fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2658–2669, Jun. 2020.
- [5] G. Xu, M. Liu, Z. Jiang, W. Shen, and C. Huang, "Online fault diagnosis method based on transfer convolutional neural networks," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 2, pp. 509–520, Feb. 2020.
- [6] J. De Viaene, F. Verbelen, S. Derammelaere, and K. Stockman, "Energy-efficient sensorless load angle control of a BLDC motor using sinusoidal currents," *IET Electr. Power Appl.*, vol. 12, no. 9, pp. 1378–1389, Sep. 2018.
- [7] A. Mousmi, A. Abbou, and Y. E. Houm, "Binary diagnosis of Hall effect sensors in brushless DC motor drives," *IEEE Trans. Power Electron.*, vol. 35, no. 4, pp. 3859–3868, Apr. 2020.
- [8] Q. Zhang and M. Feng, "Fast fault diagnosis method for Hall sensors in brushless DC motor drives," *IEEE Trans. Power Electron.*, vol. 34, no. 3, pp. 2585–2596, Mar. 2019.
- [9] S. Grubic, J. M. Aller, B. Lu, and T. G. Habetler, "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.*, vol. 55, no. 12, pp. 4127–4136, Dec. 2008.
- [10] M.-A. Shamsi-Nejad, B. Nahid-Mobarakeh, S. Pierfederici, and F. Meibody-Tabar, "Fault tolerant and minimum loss control of double-star synchronous machines under open phase conditions," *IEEE Trans. Ind. Electron.*, vol. 55, no. 5, pp. 1956–1965, May 2008.
- [11] F. Zidani, D. Diallo, M. E. H. Benbouzid, and R. Nait-Said, "A fuzzy-based approach for the diagnosis of fault modes in a voltage-fed PWM inverter induction motor drive," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, pp. 586–593, Feb. 2008.
- [12] A. Picot, Z. Obeid, J. Régner, S. Poignant, O. Darnis, and P. Maussion, "Statistic-based spectral indicator for bearing fault detection in permanent-magnet synchronous machines using the stator current," *Mech. Syst. Signal Process.*, vol. 46, no. 2, pp. 424–441, Jun. 2014.
- [13] S. Rajagopalan, J. M. Aller, J. A. Restrepo, T. G. Habetler, and R. G. Harley, "Analytic-wavelet-ridge-based detection of dynamic eccentricity in brushless direct current (BLDC) motors functioning under dynamic operating conditions," *IEEE Trans. Ind. Electron.*, vol. 54, no. 3, pp. 1410–1419, Jun. 2007.
- [14] W. Le Roux, R. G. Harley, and T. G. Habetler, "Detecting rotor faults in low power permanent magnet synchronous machines," *IEEE Trans. Power Electron.*, vol. 22, no. 1, pp. 322–328, Jan. 2007.
- [15] M. Kang, J. Kim, and J.-M. Kim, "High-performance and energy-efficient fault diagnosis using effective envelope analysis and denoising on a general-purpose graphics processing unit," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2763–2776, May 2015.
- [16] M. Kang, J. Kim, J.-M. Kim, A. C. C. Tan, E. Y. Kim, and B.-K. Choi, "Reliable fault diagnosis for low-speed bearings using individually trained support vector machines with kernel discriminative feature analysis," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2786–2797, May 2015.
- [17] B. M. Ebrahimi, J. Faiz, and M. J. Roshtkhari, "Static-, dynamic-, and mixed-eccentricity fault diagnoses in permanent-magnet synchronous motors," *IEEE Trans. Ind. Electron.*, vol. 56, no. 11, pp. 4727–4739, Nov. 2009.
- [18] J. Zhang, J. Zhao, D. Zhou, and C. Huang, "High-performance fault diagnosis in PWM voltage-source inverters for vector-controlled induction motor drives," *IEEE Trans. Power Electron.*, vol. 29, no. 11, pp. 6087–6099, Nov. 2014.
- [19] S. Nandi, H. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electric motors—A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Dec. 2005.
- [20] A. L. Martinez-Herrera, L. M. Ledesma-Carrillo, M. Lopez-Ramirez, S. Salazar-Colores, E. Cabal-Yepez, and A. Garcia-Perez, "Gabor and the Wigner-Ville transforms for broken rotor bars detection in induction motors," in *Proc. Int. Conf. Electron., Commun. Comput. (CONIELECOMP)*, Feb. 2014, pp. 83–87.
- [21] Q. Song, J. Ni, and G. Wang, "A fast clustering-based feature subset selection algorithm for high-dimensional data," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 1, pp. 1–14, Jan. 2013.
- [22] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," *IEEE Trans. Neural Netw.*, vol. 20, no. 2, pp. 189–201, Feb. 2009.
- [23] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.
- [24] F. Lin, C. Chen, and J. Lin, "Detection of mechanical resonance frequencies for interior permanent magnet synchronous motor servo drives based on wavelet multiresolution filter," *J. Eng.*, vol. 2020, no. 10, pp. 827–833, Oct. 2020.
- [25] X.-Y. Liu, Y. Liang, S. Wang, Z.-Y. Yang, and H.-S. Ye, "A hybrid genetic algorithm with wrapper-embedded approaches for feature selection," *IEEE Access*, vol. 6, pp. 22863–22874, 2018.
- [26] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 4, pp. 491–502, Apr. 2005.
- [27] C. Liu, C. Wu, and L. Jiang, "Evolutionary clustering framework based on distance matrix for arbitrary-shaped data sets," *IET Signal Process.*, vol. 10, no. 5, pp. 478–485, 2016.
- [28] Q. Fan and X. Yan, "Self-adaptive differential evolution algorithm with zoning evolution of control parameters and adaptive mutation strategies," *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 219–232, Jan. 2016.
- [29] F. Filippetti, G. Franceschini, C. Tassoni, and P. Vas, "Recent developments of induction motor drives fault diagnosis using AI techniques," *IEEE Trans. Ind. Electron.*, vol. 47, no. 5, pp. 994–1004, Oct. 2000.
- [30] J.-W. Lin, C.-T. Chao, and J.-S. Chiou, "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*, vol. 6, pp. 52582–52597, 2018.
- [31] A. A. Abohagar and M. W. Mustafa, "Back propagation neural network aided wavelet transform for high impedance fault detection and faulty phase selection," in *Proc. IEEE Int. Conf. Power Energy (PECon)*, Dec. 2012, pp. 790–795.
- [32] F. Gaxiola, P. Melin, F. Valdez, and O. Castillo, "Backpropagation learning method with interval type-2 fuzzy weights in neural networks," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Aug. 2013, pp. 1–6.
- [33] J. C. Montesdeoca, J. C. Zambrano-Abad, J. A. Morales-Garcia, and R. S. Avila-Campoverde, "Virtual speed sensor for DC motor using back-propagation artificial neural networks," in *Proc. IEEE Int. Autumn Meeting Power, Electron. Comput. (ROPEC)*, Nov. 2014, pp. 1–5.

- [34] H. Tang, Z. Liao, P. Chen, D. Zuo, and S. Yi, "A novel convolutional neural network for low-speed structural fault diagnosis under different operating condition and its understanding via visualization," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021.
- [35] Y. Gritli, M. Mengoni, C. Rossi, F. Filippetti, and D. Casadei, "New quantitative rotor fault evaluation in wound rotor induction machine drives under time-varying conditions," in *Proc. 9th IEEE Int. Symp. Diagnostics Electr. Mach., Power Electron. Drives (SDEMPED)*, Aug. 2013, pp. 57–63.
- [36] S. Choi, E. Pazouki, J. Baek, and H. R. Bahrami, "Iterative condition monitoring and fault diagnosis scheme of electric motor for harsh industrial application," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1760–1769, Mar. 2015.
- [37] L. Stankovic, T. Thayaparan, and M. Dakovic, "Signal decomposition by using the S-method with application to the analysis of HF radar signals in sea-clutter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4332–4342, Nov. 2006.
- [38] L. Zao and R. Coelho, "On the estimation of fundamental frequency from nonstationary noisy speech signals based on the Hilbert–Huang transform," *IEEE Signal Process. Lett.*, vol. 25, no. 2, pp. 248–252, Feb. 2018.
- [39] K. Price and R. Storn, "Differential evolution: A simple evolution strategy for fast optimization," *Dr. Dobb's J.*, vol. 22, no. 4, pp. 18–24, Dec. 1997.
- [40] S. Strasser, J. Sheppard, N. Fortier, and R. Goodman, "Factored evolutionary algorithms," *IEEE Trans. Evol. Comput.*, vol. 21, no. 2, pp. 281–293, Apr. 2017.
- [41] K. T. Hu, Z. Liu, K. Huang, C. Dai, and S. Gao, "Improved differential evolution algorithm of model-based diagnosis in traction substation fault diagnosis of high-speed railway," *IET Elect. Syst. Transp.*, vol. 6, no. 3, pp. 163–169, Sep. 2016.
- [42] M. Secmen and M. F. Tasgetiren, "Ensemble of differential evolution algorithms for electromagnetic target recognition problem," *IET Radar, Sonar Navigat.*, vol. 7, no. 7, pp. 780–788, Aug. 2013.
- [43] P. Sonego, A. Kocsor, and S. Pongor, "ROC analysis: Applications to the classification of biological sequences and 3D structures," *Briefings Bioinf.*, vol. 9, no. 3, pp. 198–209, Jan. 2008.
- [44] *UCI Machine Learning Repository*. Accessed: Sep. 5, 2019. [Online]. Available: <http://archive.ics.uci.edu/ml>
- [45] D. Lin, L. Sun, K.-A. Toh, J. B. Zhang, and Z. Lin, "Twin SVM with a reject option through ROC curve," *J. Franklin Inst.*, vol. 355, no. 4, pp. 1710–1732, Mar. 2018.
- [46] J. Hernández-Orallo, "ROC curves for regression," *Pattern Recognit.*, vol. 46, no. 12, pp. 3395–3411, Dec. 2013.
- [47] R. C. Luo, H.-C. Lin, and Y.-T. Hsu, "CNN based reliable classification of household chores objects for service robotics applications," in *Proc. IEEE 17th Int. Conf. Ind. Informat. (INDIN)*, Jul. 2019, pp. 547–552.
- [48] R. Guo, X. Shen, and X. Zhang, "3D ROC histogram: A new ROC analysis tool incorporating information on instances," *IEEE Access*, vol. 7, pp. 183396–183404, 2019.



CHUN-YAO LEE (Member, IEEE) received the Ph.D. degree from the Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, in 2007. From 1998 to 2008, he was a Distribution System Designer with the Engineering Division, Taipei City Government, and CECI Engineering Consultants, Inc., Taiwan. From 2004 to 2006, he was a Visiting Scholar with the Department of Electrical Engineering, University of Washington, Seattle, WA, USA, sponsored by the Ministry of Science and Technology, Taiwan. Since 2008, he has been a Faculty Member of electrical engineering with Chung Yuan Christian University, Taoyuan, Taiwan, where he is currently a Full Professor. His research interests include power distribution, optimization algorithms, and damage diagnosis. He received the National Excellent Teachers Award in Taiwan, in 2020.



CHEN-HSU HUNG was born in Taichung, Taiwan, in 1998. He received the B.S. degree, in 2020. He is currently pursuing the M.S. degree in electrical engineering with Chung Yuan Christian University, Taoyuan, Taiwan. His research interests include fault detection and optimization algorithms.



TRUONG-AN LE received the M.S. degree in electronic engineering from the University of Transport and Communications, Vietnam, in 2012, and the Ph.D. degree from the Department of Electrical Engineering, Chung Yuan Christian University, Taiwan, in 2022. He is currently a Faculty Member with the Institute of Engineering and Technology, Thu Dau Mot University, Binh Duong, Vietnam. His research interests include fault detection, optimization algorithms, machine learning, and deep learning.

...