This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications

Machine Learning based Fault Diagnosis for Single- and Multi-Faults in Induction Motors Using Measured Stator Currents and Vibration Signals

Mohammad Zawad Ali
Student Member, IEEE
Memorial University of Newfoundland
Department of Electrical and Computer Engineering
St John's, NL, Canada
mzali@mun.ca

Md Nasmus Sakib Khan Shabbir
Student Member, IEEE
Memorial University of Newfoundland
Department of Electrical and Computer Engineering
St John's, NL, Canada
mnskshabbir@mun.ca

Xiaodong Liang
Senior Member, IEEE
Memorial University of Newfoundland
Department of Electrical and Computer Engineering
St John's, NL, Canada
xliang@mun.ca

Yu Zhang
Memorial University of Newfoundland
Department of Computer Science
St John's, NL, Canada
yzhang17@mun.ca

Ting Hu
Memorial University of Newfoundland
Department of Computer Science
St John's, NL, Canada
ting.hu@mun.ca

***Abstract* – In this paper, a practical machine learning based fault diagnosis method is proposed for induction motors using experimental data. Various single- and multi- electrical and/or mechanical faults are applied to two identical induction motors in lab experiments. Stator currents and vibration signals of the motors are measured simultaneously during experiments and are used in developing the fault diagnosis method. Two signal processing techniques, Matching Pursuit (MP) and Discrete Wavelet Transform (DWT), are chosen for feature extraction. Three classification algorithms, support vector machine (SVM), K-nearest neighbors (KNN), and Ensemble, with 17 different classifiers offered in MATLAB Classification Learner toolbox are used in the study to evaluate the performance and suitability of different classifiers for induction motor fault diagnosis. It is found that five classifiers (Fine Gaussian SVM, Fine KNN, Weighted KNN, Bagged trees, and Subspace KNN) can provide near 100% classification accuracy for all faults applied to each motor, but the remaining 12 classifiers do not perform well. A novel curve fitting technique is developed to calculate features for the motors that stator currents or vibration signals under certain loadings are not tested for a particular fault. The proposed fault diagnosis method can accurately detect single- or multi- electrical and mechanical faults in induction motors.**

***Index Terms*—Discrete wavelet transform, fault diagnosis, induction motors, machine learning, matching pursuit.**

I. INTRODUCTION

Fault diagnosis of induction motors is critical to maintain uninterrupted operation of industrial processes. In the literature, there are three streams of research on fault diagnosis for induction motors: 1) signature extraction based approaches; 2) model-based approaches; and 3) knowledge-based approaches. The signature extraction based approaches are achieved by surveying fault signatures in time and/or frequency domain. Current, voltage, power, vibration, temperature, and acoustic emission can serve as monitoring signals. Signatures extracted from the recorded monitoring signals are used to detect faults. Motor Current Signature Analysis (MSCA), a well-known spectral analysis method, is one of the most popular techniques for online monitoring induction motors in industrial environments. The MCSA can remotely monitor the stator current through the motor control center, and is most successful in detecting broken rotor bars or end rings faults. However, the false fault indication is a common issue with MSCA that needs to be improved [1]. The model-based approaches rely on mathematical modeling to predict behaviors of induction motors under fault conditions. Although model-based approaches can provide warnings and estimate incipient faults, its accuracy is largely dependent on explicit motor models, which may not be always available. The knowledge-based approaches, on the other hand, do not need a trigger threshold, machine models, motor or load characteristics. Knowledge-based approaches use machine learning techniques for on-line and off-line applications. Artificial intelligence methods have been applied for fault diagnosis in very complex time-varying and non-linear systems. With continuous advancement of machine learning algorithms, the knowledge-based approach emerges as a promising research direction for induction motor fault diagnosis with great industrial application potential.

During past two decades, the most reported machine learning methods for fault diagnosis of induction motors are the artificial neural network (ANN) or hybrid ANN combined with other techniques [2]-[15]. As one appealing feature of

This project was funded in part by IEEE Foundation through IEEE IAS Myron Zucker Faculty-Student Grant.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications

ANN that can be used for on-line applications, many of the proposed ANN methods are for on-line fault diagnosis of induction motors [2]-[5]. The hybrid ANN methods include: Park's vector–neural networks approach [5], analytical redundancy method based neural network modeling [7], statistical and neural network approaches [8][9], convolutional discriminative feature learning method [10]. One of the most popular hybrid ANN methods is combining ANN with Fuzzy logic, which can provide accurate fault detection with heuristic interpretation [11]-[15].

Several other machine learning approaches are employed in [16]-[20]. The immunological principles are applied for induction motor fault detection in [16]. A pattern recognition approach associated with Kalman interpolator/extrapolator is proposed in [17]. An integrated class-imbalanced learning scheme for diagnosing bearing defects is reported in [18]. A sparse deep learning method proposed in [19] can overcome overfitting risk of deep networks. In [20], signal processing and machine-learning techniques are combined for bearing fault detection, a novel hybrid approach based on Optimized Stationary Wavelet Packet Transform (Op-SWPT) for feature extraction and Artificial Immune System (AIS) nested within Support Vectors Machines (SVM) for fault classification is proposed. Investigations conducted using multiple machine learning algorithms are reported in [21][22].

Among machine learning based fault diagnosis methods, stator current is the most widely used signal, either alone or combined with other signals. The stator current alone is reported in [2]-[5],[8][15][16],[20]-[22]; vibration signal alone is reported in [6][7][9][10]; stator current and rotor speed combined is reported in [11][12]; stator current, speed, load and friction combined is reported in [13]; stator current, speed, winding temperature, bearing temperature and noise combined is reported in [14]; and stator current and voltage combined is reported in [17]. It appears that stator currents and vibration signals are two dominant signals used in induction motor fault diagnosis by signature extraction based approaches [1] and machine learning based approaches. However, no quantitative comparison is reported in the literature between stator currents and vibration signals for their fault diagnosis accuracy.

Despite various reported machine learning based fault diagnosis methods for induction motors, these methods have not been as widely used in real life as other techniques such as MSCA. Practical approaches need to be developed in industrial applications to take advantage of advanced and intelligent nature of machine learning.

To fill in this research gap, a practical machine learning based approach for induction motor fault diagnosis is proposed using experimental data in this paper. Experiments were conducted on two identical induction motors under healthy, single- and multi-fault conditions. A total of six motor loadings were tested for each healthy or faulty case. Stator currents and vibration signals of the motors were measured simultaneously in each testing.

Machine learning relies on features extracted from measurement signals [23]. In this paper, two signal processing techniques are adopted for feature extraction: Discrete Wavelet Transform (DWT) and Matching Pursuit (MP). Most DWT applications are for spectral analysis through the MSCA and threshold decision [24], where start-up or transient motor currents are analyzed [25][26]. However, DWT is rarely used for feature extraction [23]. Matching Pursuit was invented and firstly reported in [27] by Mallat and Zhang in 1993, and only a few papers are found so far implementing MP for induction motor fault diagnosis [28]-[31]. In this paper, the suitability of MP and DWT for feature extraction for induction motor fault diagnosis is evaluated.

The major contribution of the paper includes: 1) An effective machine learning based fault diagnosis method is proposed for single- and multi-fault diagnosis of induction motors using experimental data; 2) Both measured stator currents and vibration data are used to detect faults, and their quantitative comparison on the fault classification accuracy for the same groups of faults is demonstrated for the first time; 3) MP and DWT as signal processing methods are evaluated for feature extraction; 4) Three classification algorithms, SVM, K-nearest neighbors (KNN), and Ensemble, are evaluated with 17 different classifiers offered in MATLAB Classification Learner toolbox, and the effectiveness of chosen classifiers is compared; 5) Experiments were only conducted for six motor loadings in this study, different motor loadings between training and testing processes can deeply influence the fault diagnosis, to avoid this drawback, curve fitting equations are developed in this paper to calculate unknown features for any untested motor loadings.

The paper is arranged as follows: in Section II, the machine learning based fault diagnosis approach using experimental data is proposed; Experimental set-up is provided in Section III; in Section IV, signal processing using MP and DWT is conducted using measured stator current and vibration data, eight features are extracted through MP or DWT processing; classification accuracies using different classifiers are demonstrated in Section V; In Section VI, curve fitting equations are developed to calculate unknown features vs. motor loadings; conclusions are drawn in Section VII.

II. The Proposed Machine Learning Based Fault Diagnosis Approach

In this paper, an effective machine learning based fault diagnosis approach using experimental data is proposed. The main idea is illustrated in Fig. 1.

Six steps are needed to implement this method: 1) Conduct experiments for an induction motor under healthy, single- and multi-fault conditions; 2) Record stator current and vibration data simultaneously, where vibration sensors and a power quality analyzer are required to take measurements; 3) Choose suitable signal processing methods, such as MP and DWT, for features extraction; 4) Extract features for machine learning;

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications

5) Conduct classification for electrical and mechanical faults using chosen classifiers; and 6) Develop curve fitting equations to calculate features vs. motor loadings.

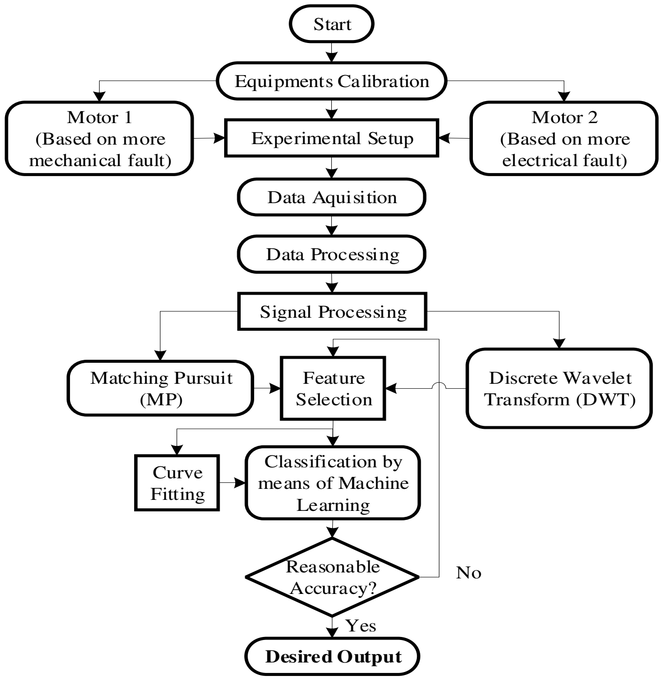


Fig. 1. The flow chart of the proposed method.

III. Experimental Set-Up

In this paper, 4-pole, 0.25 HP, 208-230/460V, 1725 rpm rated squirrel-cage induction motors (Model LEESON 101649) are purchased for experiments in the lab. Two identical motors named as “Motor 1” and “Motor 2”, which are treated as sister units, are used. Motor 1 is mainly tested for mechanical faults, and Motor 2 for electrical faults. The healthy, single- and multi-fault conditions are applied to the two motors as shown in Fig. 2.

Motor 1 was tested for the healthy condition (H), plus two single faults and three multi-faults: (a) an unbalance shaft rotation (UNB); (b) a bearing fault (BF); (c) a combined BF and UNB fault; (d) a combined BF and one broken rotor bar (BRB) faults; and (e) a combined BF, UNB, and unbalance voltage (UV) condition from the three-phase power supply.

Similarly, Motor 2 was tested for the healthy condition (H), plus four single faults and one multi-fault: (a) a UV from three-phase power supply; (b) one BRB fault; (c) two BRB fault; (e) three BRB fault; and (f) a combined UV and three BRB fault.

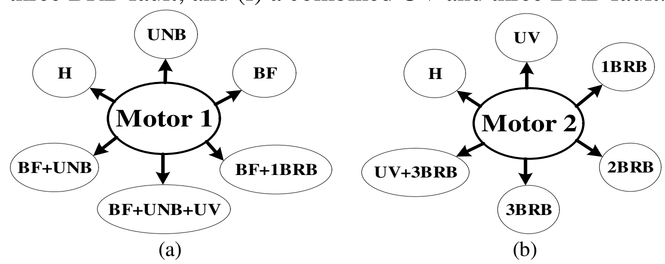


Fig. 2. Experimental plan of the applied faults: (a) Motor 1; (b) Motor 2.

In the experimental test bench (Fig. 3), an induction motor is connected directly to a three-phase power supply, and a dynamometer coupled to the motor shaft through a belt pulley serves as the load. Motor loadings can be adjusted by the dynamometer’s control knob. Under full load, the torque is 7 pound force inch (lbf-in) at the rated speed.

As shown in Fig. 4, an eight-channel power quality analyzer, PQPro by CANDURA instrument, is used to monitor and record three-phase currents. The vibration is measured by a tri-axial accelerometer (Model 356A32) with a four-channel sensor signal conditioner (Model 482C05). The accelerometer is mounted on the top of the motor near the face end, vibration at the axial (x-axis), vertical (y- axis) and horizontal (z-axis) directions is measured. A 4-channel oscilloscope is patched

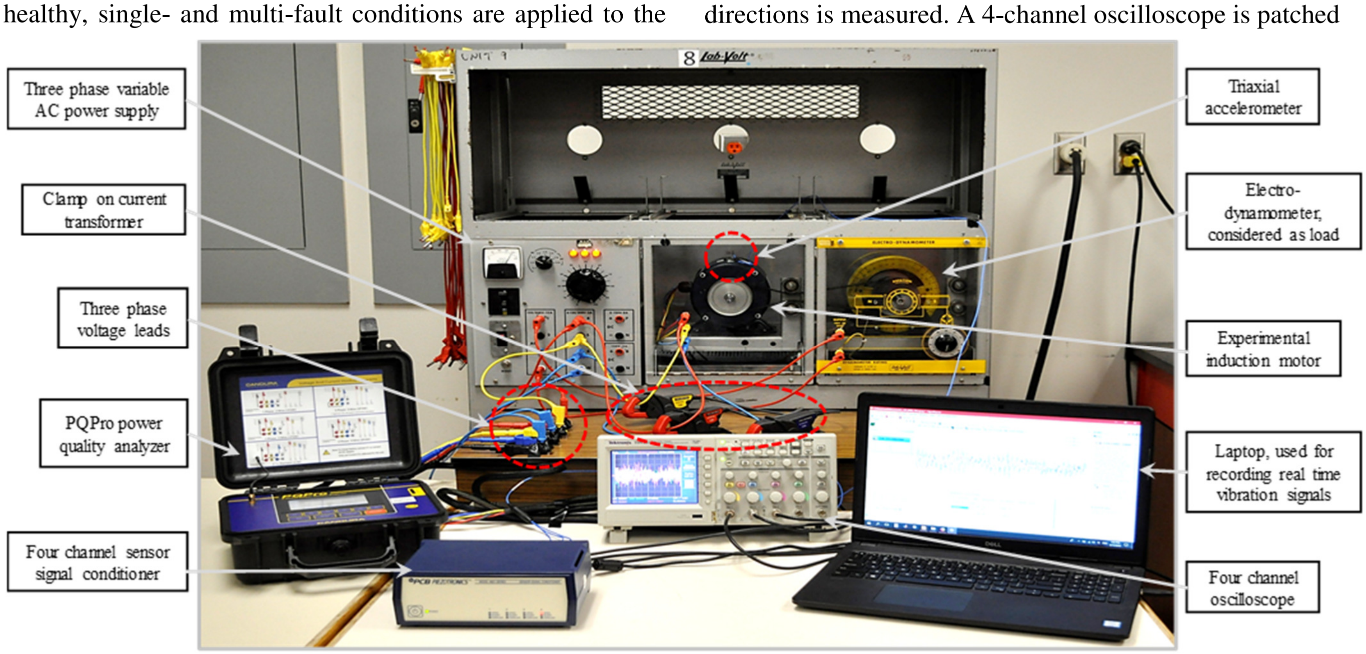


Fig.3. Experimental test bench used in this study.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications

between the sensor signal conditioner and the computer for vibration data acquisition. The sampling frequency for vibration measurements is 1.5 kHz. In each test, three phase stator currents (I_1, I_2, and I_3) and vibration at x-, y-, and z-axis during the start-up and steady-state conditions are recorded simultaneously for two minutes. A single- or multi-fault creates unbalance inside the motor, which will be reflected in stator currents and vibration signals.

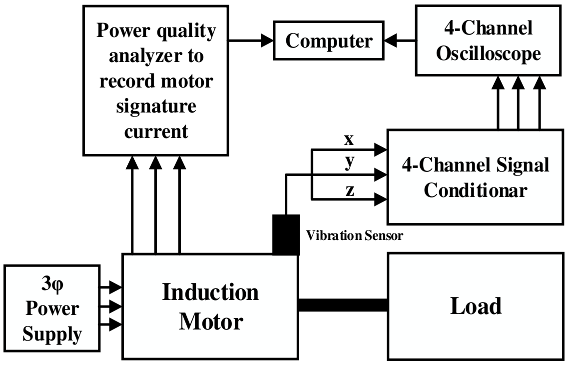
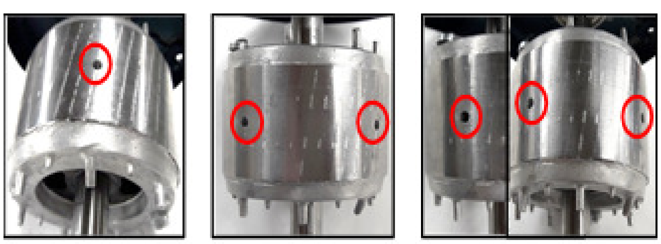
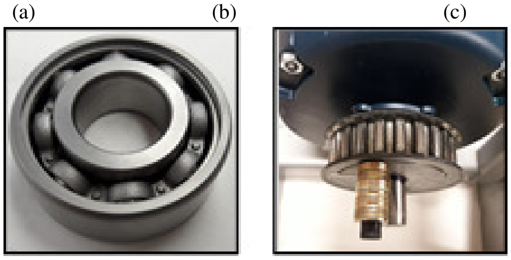


Fig. 4. Experimental schematic diagram for the system set-up.

In experiments, a BRB fault was realized by drilling a hole of a 4.2 mm diameter and 18 mm depth in the rotor bar. One hole was drilled for one BRB fault (Fig. 5 (a)); two and three holes with 90° separation were drilled for two and three BRB faults, respectively (Figs. 5 (b) and (c)). The bearing fault was the general roughness type, realized by a sand blasting process, outer and inner raceway of the bearing became very rough as shown in Fig. 5 (d). The UNB is due to uneven mechanical load distribution causing unbalanced shaft rotation. The UNB was created by adding extra weight on part of the pulley (Fig. 5 (e)). An UV condition was produced by adding an extra resistance at the second phase of the power supply for the motor. Six different loadings (10%, 30%, 50%, 70%, 85% and 100%) of the motors were tested for each fault.



(a) (b) (c)



(d) (e)

Fig. 5. Implementation of different faults in the experimental test bench: (a) 1 BRB, (b) 2 BRB, (c) 3 BRB, (d) bearing fault – general roughness type, and (e) UNB condition.

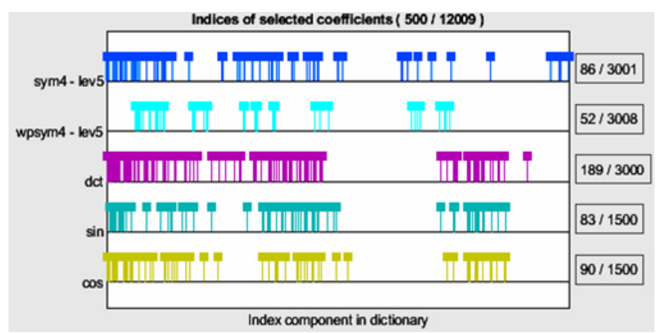
IV. Signal Processing for Feature Extraction

In this paper, two signal processing algorithms, MP and DWT, are adopted for feature extraction through MATLAB Wavelet toolbox.

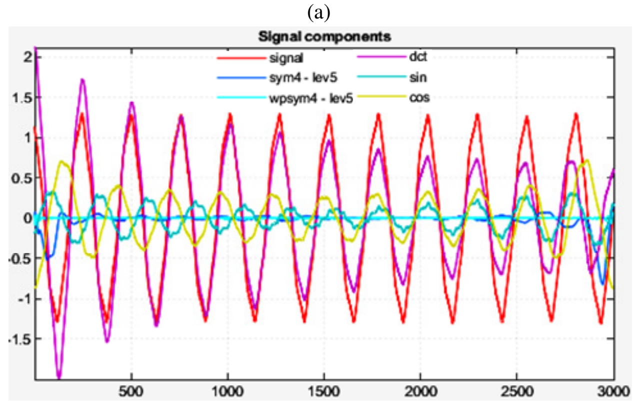
A. *Matching Pursuit*

Matching Pursuit decomposes a signal into a linear expansion of waveforms (atoms) that are selected from a redundant dictionary of functions to best match original signal [27]-[30]. To simplify the problem, only the measured stator current at the second phase (I_2) and vibration at z-axis are used for signal processing by the orthogonal matching pursuit (OMP) technique.

As an example, MP processing results for Motor 2 with a 1 BRB fault at 100% loading are shown in Fig. 6 using the current I_2 and Fig. 7 using the z-axis vibration signal. In these figures, "3000" at the x-axis means 3000 sample points. In Figs. 6 (a) and 7 (a), five signal components are chosen from the MP dictionary: 1) "sym4-lev5", the Daubechies least-asymmetric wavelet with 4 vanishing moments at the 5th level; 2) "wpsym4-lev5", the Daubechies least-asymmetric wavelet packet with 4 vanishing moments at 5th level; 3) "dct", the discrete cosine transform-II basis; 4) "sin", the Sine subdictionary; and 5) "cos", Cosine subdictionary [32]. The dct and cos components are dominant in Fig. 6 (a), and the dct and sym4-lev5 components are dominant in Fig. 7 (a). By OMP processing, the approximated signals in Figs. 6 (c) and 7 (c) are obtained by setting the "maximum relative error" of "L1 Norm" equal to 0.01%, and the "maximum iterations" equal to 100 in the MATLAB Wavelet toolbox. With the same procedure, all measured current and vibration signals under healthy and faulty conditions for the two motors are analyzed.

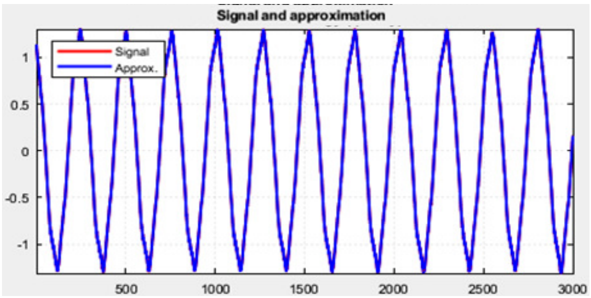


(a)



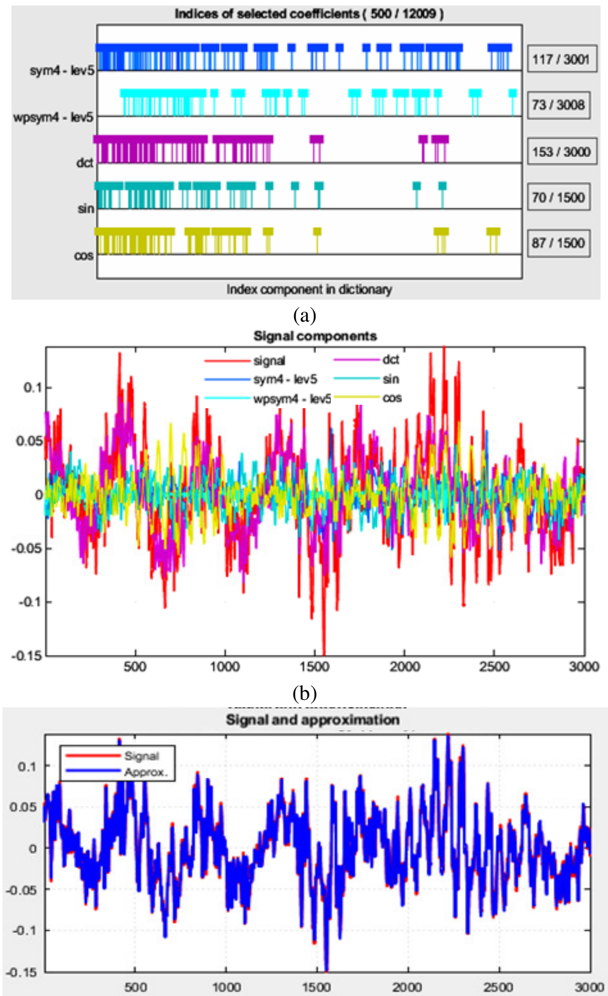
(b)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications



(c)

Fig. 6. The stator current I_2 for Motor 2 using MP (1 BRB fault, 100% loading): (a) indices of selected coefficients; (b) original signal and signal components; (c) signal and its approximation.



(a)

(b)

(c)

Fig. 7. The z-axis vibration signal for Motor 2 using MP (1 BRB fault, 100% loading): (a) indices of selected coefficients; (b) original signal and signal components; (c) signal and its approximation.

Eight statistical features are determined using the OMP as follows: mean, median, standard deviation, median absolute deviation, mean absolute deviation, L1 norm, L2 norm, and the maximum norm as tabulated in Table I [33][34]. Table II shows a sample of features obtained using the current I_2 for Motor 2 with a 1BRB fault at 100% loading. Every set of eight features, such as S1 in the first row of Table II, is obtained by taking 3000 sample points from the current I_2 and processed by the OMP. Other sets of features (from S_2 to S_7) are determined by taking sample points in a similar way.

Fig. 8 shows one feature, Mean, for Motors 1 and 2 processed by the current I_2 vs. motor loadings and different types of faults. Other features show similar patterns.

TABLE I
STATISTICAL FEATURES [33][34]

Features	Formations
Mean	$\mu_X = \frac{1}{N}\sum_{i=1}^{N} x_i$, where x_i is the ith sampled measurement point, i = 1, 2, 3, …, N for N observations.
Median	med $= \frac{1}{2}(x_{(\lfloor(N+1)/2\rfloor)} + x_{(\lfloor N/2\rfloor+1)})$
Standard Deviation (Std. Dev.)	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)^2}$, where μ_x is the mean.
Median Absolute Deviation	Median_AD = $median(\lvert x_i - meidan(X)\rvert)$
Mean Absolute Deviation	Mean_AD = $\frac{1}{N}\sum_{i=1}^{N}\lvert x_i - \mu_x\rvert$
L1 norm	$\lVert L\rVert_1 = \sum_{i=1}^{N}\lvert x_i\rvert$, the sum of absolute values of its components, also known as one-norm, or mean norm
L2 norm	$\lVert L\rVert_2 = \sqrt{\sum_{i=1}^{N}\lvert x_i\rvert^2}$, the square root of the sum of the squares of absolute values of its components, also known as two-norm, or mean-square norm.
Maximum norm (Max norm)	$\lVert L\rVert_\infty = \max\{\lvert x_i\rvert : i = 1, 2, \ldots, n\}$, the maximum of absolute values of its components, also known as infinity norm, or uniform norm.

B. Discrete Wavelet Transform

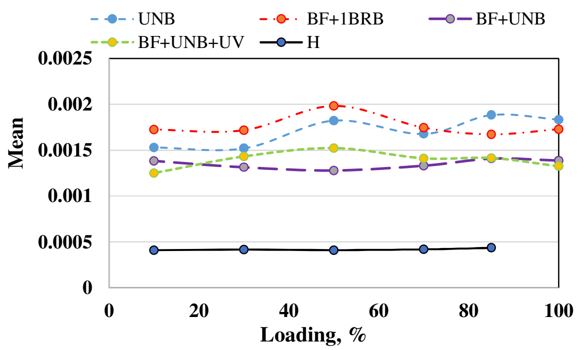
Wavelet transform defines a signal consisting of regions of different frequency components. It can decompose a signal into wavelets confined by both time and frequency [25][35]. In this paper, motor stator currents and vibration signals are analyzed using the DWT analysis. The wavelet db4 is selected as the mother wavelet under consideration of the 6th level decomposition. db4 is from the Daubechies family with four vanishing moment. To demonstrate the DWT processing results, the stator current I_2 and z-axis vibration signals for Motor 2 with a 1 BRB fault at 100% motor loading are analyzed as shown in Figs. 9 and 10, respectively.

Similar to MP, the aim of the DWT processing is to extract statistical features of the original signal after the signal decomposition. Through the DWT analysis, eight features defined in Table I are also determined. Table III shows a sample of eight features processed using the stator current I_2 for Motor 2 with a 1BRB fault at 100% loading.

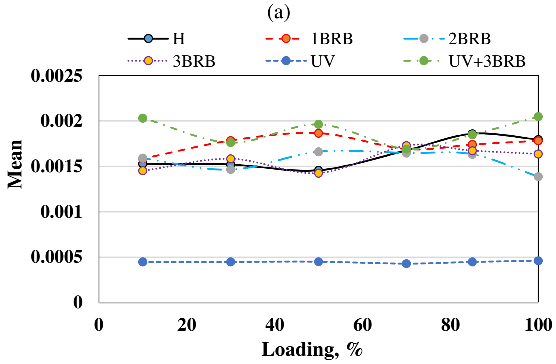
TABLE II
A SAMPLE OF FEATURES USING STATOR CURRENT I_2 PROCESSED BY OMP (MOTOR 2, 1 BRB, 100% LOADING)

Features	Mean	Median	Std. Dev.	Median Absolute Dev.	Mean Absolute Dev.	L1 norm	L2 norm	Max norm
s1	0.001783	0.001462	0.001397	0.0008932	0.0011080	5.349	0.1241	0.008743

s2	0.001624	0.001341	0.001261	0.0007733	0.0009930	4.873	0.1126	0.007977
s3	0.001676	0.001400	0.001284	0.0008274	0.0010160	5.027	0.1156	0.009878
s4	0.001545	0.001285	0.001205	0.0007696	0.0009574	4.634	0.1073	0.006766
s5	0.001770	0.001458	0.001351	0.0008747	0.0010750	5.310	0.1220	0.009370
s6	0.001583	0.001331	0.001223	0.0008150	0.0009799	4.750	0.1096	0.007019
s7	0.001712	0.001460	0.001305	0.0008588	0.0010350	5.135	0.1179	0.007477



(a)



(b)

Fig. 8. One feature, Mean, vs. motor loadings and different types of faults processed by OMP using the stator current I_2: (a) Motor 1; (b) Motor 2.

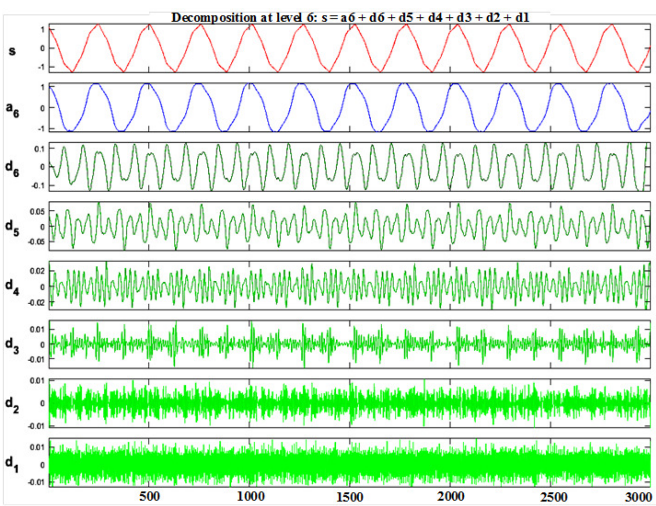


Fig. 9. The processed one phase stator current signal I_2 using DWT for Motor 2 under a 1 BRB fault and 100% loading condition.

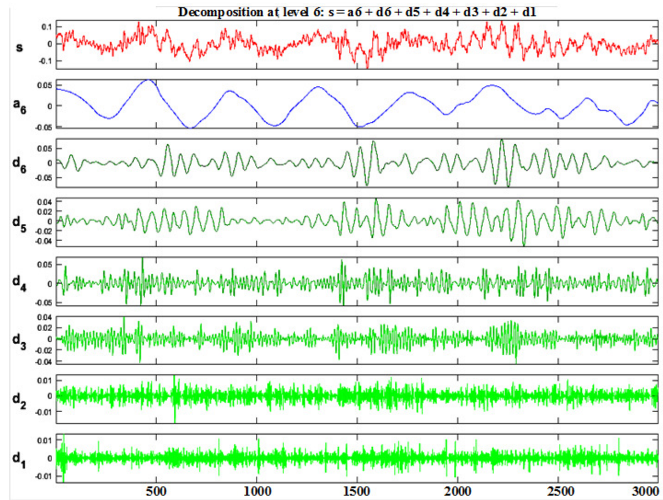


Fig. 10. The processed z-axis vibration signal using DWT for Motor 2 under a 1 BRB fault and 100% loading condition.

V. Machine Learning Results

Several classification algorithms are available in the MATLAB Classification Learner Toolbox. In this paper, three algorithms, SVM, KNN, and ensemble, are selected with 17 different classifiers. Their performance and suitability for induction motor fault diagnosis are evaluated.

A. Classification Algorithms

SVM is a commonly used machine learning method for data classification and regression based on statistical learnings and structural risk minimization [38]. It generally classifies a dataset into two classes, positive and negative. A statistical learning theory based algorithm is used to train the data set, which is known as support vector. It provides information about the classification and builds the hyperplane. The hyperplane maximizes the margin of separation between positive and negative classes [39]. SVM is suitable for a dataset where separable and non-separable data profile are present. The soft margin (hyperplane), which is the smallest distance in the architecture for separable and non-separable data set, is used to distinguish data points. Kernel functions are used for nonlinear transformation. A kernel function converts a nonlinearly separable object into linearly separable by mapping them in a higher dimensional feature space [23]. The common types of kernel functions include linear kernel, polynomial kernel, Gaussian radial basis function (RBF) kernel as shown in Table IV [40][41].

TABLE III
A SAMPLE OF FEATURES FOR MACHINE LEARNING USING ONE PHASE STATOR CURRENT I_2 PROCESSED BY DWT (MOTOR 2, 1 BRB, 100% LOADING)

Features	Mean	Median	Std. Dev.	Median Absolute Dev.	Mean Absolute Dev.	L1 norm	L2 norm	Max norm
s1	-0.021220	-0.040460	0.8473	0.8354	0.7623	2288	46.42	1.307
s2	-0.025300	-0.042620	0.8459	0.8357	0.7602	2282	46.34	1.309

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications

s3	-0.022740	-0.043430	0.8445	0.8314	0.7591	2278	46.26	1.308
s4	-0.020420	0.039110	0.8474	0.8419	0.7626	2289	46.42	1.316
s5	-0.013450	-0.034260	0.8522	0.8473	0.7686	2306	46.67	1.303
s6	-0.004517	-0.007013	0.8570	0.8583	0.7733	2320	46.93	1.309
s7	0.006022	0.013220	0.8558	0.8543	0.7721	2317	46.87	1.307

TABLE IV
COMMON SVM KERNEL FUNCTIONS [40][41]

Kernel name	Kernel function formula	Description
Linear kernel	$k(x,y) = x^T y + c$	Linear kernel is the simplest kernel function. It is given by the inner product (x,y) plus an optional constant **c.**
Polynomial Kernel	$k(x,y) = (\alpha x^T y + c)^d$ Where, adjustable parameters are the slope alpha, the constant term is c and the polynomial degree is d.	Polynomial kernel is a non-stationary kernel, well suited for problems where all the training data is normalized. The most common degree is d = 2 (quadratic) and d = 3 (cubic), since larger degree tends to overfit on machine learning problems.
Gaussian Kernel or Radial Basis Function (RBF)	$k(x,y) = exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$ or $k(x,y) = exp(-\gamma\|x-y\|^2)$ Where, $\gamma = 1/2\sigma^2$ is an adjustable parameter and $\|x-y\|$ is denoted as squared euclidean distance between the two feature vectors.	In Gaussian kernel, γ plays a major role in the performance of the kernel. If over-estimated, the exponential will behave almost linearly and the higher-dimensional projection will start to lose its non-linear power.

KNN is an instance based classification technique that classifies an unknown instance by correlating it with a known instance through a similarity function or an effective distance. It is the simplest machine learning process to classify data. In KNN, a data set is divided into a fixed number (k) of clusters. The center data point of a cluster is called centroid, which can be real or imaginary, is used to train the KNN classifier. Choosing centroid value is an iterative process. To generate an initial set of random clusters, the emanated classifier is used. Then it continue to adjust the centroid value until it becomes stable. The stable centroids are used to classify input data by transforming an anonymous dataset into a known one [42].

Ensemble is a superior classifier that combines multiple diverse single classifier to boost the prediction accuracy. Each single classifier is trained and then combined. The combined ensemble can be trained later as a single hypothesis, which is not necessarily constrained within the set of hypothesis from where it is originated. This flexibility may lead to over fitting, which is overcome in Bagged Trees where each classifier is trained in different partitions and combined through a majority voting. A weaker correlation of error of single classifiers leads to a better prediction accuracy. Therefore, diverse single classifiers are preferred for ensemble [43]-[46].

B. Classifiers Selected from the Toolbox

The MATLAB Classification Learner toolbox can train models to classify data using supervised machine learning. In this paper, three classification algorithms, SVM, KNN and Ensemble, provided in the toolbox are chosen to perform fault diagnosis. The selected 17 classifiers are listed as follows:

- SVM: linear SVM, quadratic SVM, cubic SVM, fine Gaussian SVM, medium Gaussian SVM, and coarse Gaussian SVM.
- KNN: fine KNN, medium KNN, coarse KNN, cosine KNN, cubic KNN, and weighted KNN.
- Ensemble: boosted trees, bagged trees, subspace discriminant, subspace KNN, and RUSBoosted trees.

Table V shows descriptions of each classifier used in the paper.

We performed a five-fold cross validation to protect against overfitting in this paper. The data is partitioned into five disjoint folds. For each of the five iterations, four folds were used as training samples and one fold as testing samples. Each sample in the data was used as a testing sample exactly once. The average test error is calculated over all folds. This method gives a good estimation of the predictive accuracy of the final model trained with all the data.

TABLE V
17 CLASSIFIERS FROM MATLAB CLASSIFICATION LEARNER TOOLBOX

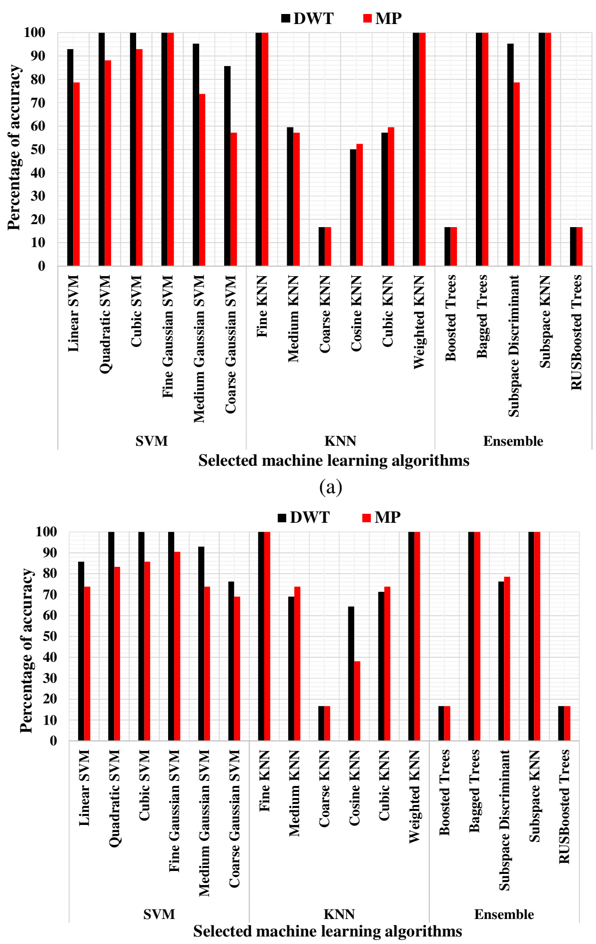
Classification algorithms	Classifier types	Classifier description from MATLAB classification learner toolbox
Support vector machines (SVM)	Linear SVM	Makes a simple linear separation between classes, using the linear kernel. The easiest SVM to interpret.
	Quadratic SVM	Uses the quadratic kernel.
	Cubic SVM	Uses the cubic kernel.
	Fine Gaussian SVM	Makes finely-detailed distinctions between classes, using the Gaussian kernel with kernel scale set to sqrt(P)/4, where P is the number of predictors.
	Medium Gaussian SVM	Makes fewer distinctions than a Fine Gaussian SVM, using the Gaussian kernel with kernel scale set to sqrt(P), where P is the number of predictors.
	Coarse Gaussian SVM	Makes coarse distinctions between the classes, using the Gaussian kernel with kernel scale set to sqrt(P)*4, where P is the number of predictors.
Nearest neighbor classifiers (KNN)	Fine KNN	Makes finely detailed distinctions between classes, with the number of neighbors set to 1.
	Medium KNN	Makes fewer distinctions than a Fine KNN, with the number of neighbors set to 10.
	Coarse KNN	Makes coarse distinctions between classes, with the number of neighbors set to 100.
	Cosine KNN	Uses a cosine distance metric, with the number of neighbors set to 10.
	Cubic KNN	Uses a cubic distance metric, with the number of neighbors set to 10.
	Weighted KNN	Uses a distance weighting, with the number of neighbors set to 10.
Ensemble classifiers	Boosted trees	This model creates an ensemble of medium decision trees using the AdaBoost algorithm. Compared to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications

		bagging, boosting algorithms use relatively little time or memory, but might need more ensemble members.
	Bagged trees	It is a bootstrap-aggregated ensemble of fine decision trees. Often very accurate, but can be slow and memory intensive for large data sets.
	Subspace discriminant	Good for many predictors, relatively fast for fitting and prediction, and low on memory usage, but the accuracy varies depending on the data. The model creates an ensemble of Discriminant classifiers using the Random Subspace algorithm.
	Subspace KNN	Good for many predictors. The model creates an ensemble of nearest-neighbor classifiers using the Random Subspace algorithm.
	RUSBoosted trees	Used for skewed data with many more observations of one class.

B. *Fault Diagnosis Results*

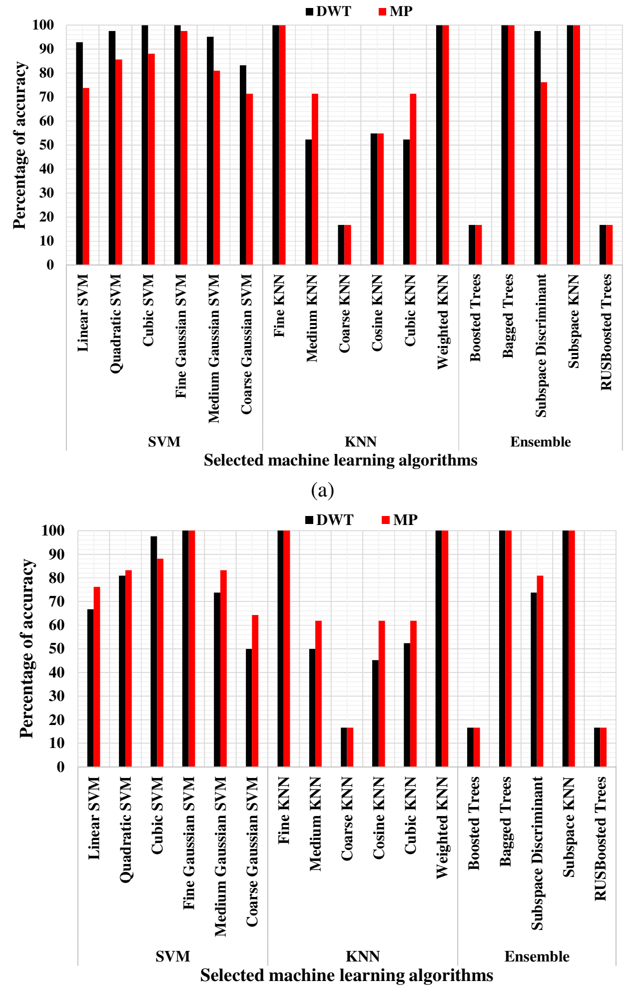
The fault diagnosis accuracies for all faults of Motors 1 and 2 at 100% loading using the current I_2 and z-axis vibration signal are shown in Figs. 11 and 12, respectively. In each graph, MP and DWT processing are compared. The data for Fig. 12 are also shown in Table VI.



(a)

(b)

Fig. 11. Classification accuracy for all faults implemented on Motor 1 at 100% loading using the selected classifiers: (a) stator current I_2; (b) z-axis vibration.



(a)

(b)

Fig. 12. Classification accuracy for all faults implemented on Motor 2 at 100% loading using the selected classifiers: (a) stator current I_2; (b) z-axis vibration.

TABLE VI
ACCURACY FOR CLASSIFICATION OF ALL FAULTS FOR MOTOR 2 AT 100% LOADING USING VARIOUS CLASSIFIERS

Classification Method	Sub-groups	Features by means of MP (% of accuracy)		Features by means of DWT (% of accuracy)	
		Current (I_2)	Vibration (z-axis)	Current (I_2)	Vibration (z-axis)
SVM	Linear	73.8	76.2	92.9	66.7
	Quadratic	85.7	83.3	97.6	81
	Cubic	88.1	88.1	100	97.6
	Fine Gaussian	**97.6**	**100**	**100**	**100**
	Medium Gaussian	81	83.3	95.2	73.8
	Coarse Gaussian	71.4	64.3	83.3	50
KNN	**Fine**	**100**	**100**	**100**	**100**
	Medium	71.4	61.9	52.4	50
	Coarse	16.7	16.7	16.7	16.7
	Cosine	54.8	61.9	54.8	45.2
	Cubic	71.4	61.9	52.4	52.4
	Weighted	**100**	**100**	**100**	**100**
Ensemble	Boosted Trees	16.7	16.7	16.7	16.7

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications

Bagged Trees	**100**	**100**	**100**	**100**
Subspace Dis-criminant	76.2	81	97.6	73.8
Subspace KNN	**100**	**100**	**100**	**100**
RUSBoosted Trees	16.7	16.7	16.7	16.7

It is found that the five classifiers, Fine Gaussian SVM, Fine KNN, Weighted KNN, Bagged trees, and Subspace KNN, return mostly 100% classification accuracy for all faults on each motor at 100% loading. The classification accuracy for other motor loadings is similar to 100% loading for these five classifiers. However, not all selected classifiers are suitable for fault diagnosis. As the worst case, the Boosted Trees and RUSBoosted Trees only have 16.7% classification accuracy.

It can be observed that DWT has better accuracy than MP for most SVM classifiers, while MP has better accuracy than DWT for most KNN algorithms. Both MP and DWT demonstrate excellent and equally strong performance, and thus, they can be used as signal processing tools to extract features for induction motor fault diagnosis.

The classifier performance is assessed using the confusion matrix and receiver operating characteristic (ROC) curve in this paper. The confusion matrix indicates how a classifier performed in each class. It is able to categorize the regions, where the classifier has performed correctly or poorly. The rows show the true class, the columns show the predicted class, and the diagonal cells show where the true class and predicted class match. If these diagonal cells are green, it means that the classifier has performed well and classified observations of this true class correctly. The accuracy in the confusion matrix is calculated as follows:

$$Accuracy = \frac{TP}{TP + FN} \qquad (1)$$

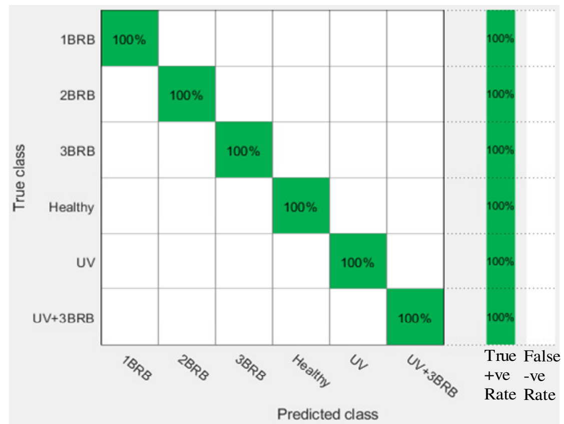
Where, TP is true positive, and FN is false negative. The ROC curve is a graphical representation of the confusion matrix. It summarizes the overall performance of a classifier over all possible thresholds, and the area under the curve (AUC) gives an insight about how confidently the classification is done. The ROC curve shows true positive rate (TPR) versus false positive rate (FPR) for a trained classifier, where TPR and FPR can be calculated as follows [47][48]:

$$\text{True positive rate} = \frac{TP}{TP + FN} = 1 - \text{False negative rate} \quad (2)$$

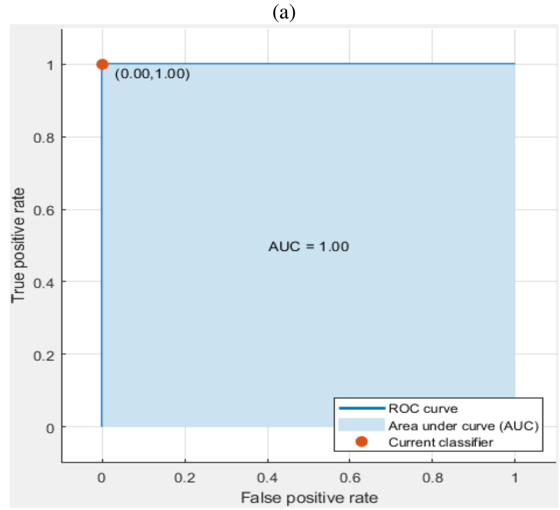
$$\text{False positive rate} = \frac{FP}{FP + TN} = 1 - \text{True negative rate} \quad (3)$$

Where, TP is true positive, FN is false negative, FP is false positive, and TN is true negative. TPR signifies how often the classifier predicts positive when the actual case is positive; FPR represents how often the classifier incorrectly predicts positive when the actual case is negative. Both TPR and FPR range from 0 to 1, and the AUC ranges from 0.5 to 1. An AUC of 1 represents a good result with no misclassified points; while an AUC of 0.5 represents that the classifier is doing no better than random guessing.

Fig. 13 shows the confusion matrix and ROC curve with 100% classification accuracy obtained by the classifier, Fine KNN, for Motor 2 at 100% loading and processed using the current I_2 signal.



(a)



(b)

Fig. 13. 100% classification accuracy obtained by Fine KNN for Motor 2 at 100% loading using the current I_2: (a) confusion matrix; (b) ROC curve.

C. Stator Current vs. Vibration Signal

In this study, both motors are tested for healthy and faulty conditions from light load to full load. Although Motor 1 has mostly mechanical faults, and Motor 2 has electrical faults, it can be observed in Figs. 11 and 12 that both stator current and vibration signals work equally well for fault diagnosis of each motor. Therefore, by the quantitative comparison through this research, it is concluded that either stator current or vibration signal can serve as the condition monitoring signal for induction motor fault diagnosis with a comparable accuracy.

In real life applications, stator currents are more readily available than vibration signals. Stator currents can be measured at the motor terminal or remotely at the motor control center; while vibration measurements require a vibration sensor attached to the motor surface, more costly and complicated, especially for motors in a harsh environment.

D. *Influence of the Number of Chosen Features*

In this study, we have chosen eight features for fault classification. It is important to evaluate the influence of the number of features on the classification accuracy. The following six cases are considered for feature selection: 1) Two features: mean and median; 2) Two features: mean and max norm; 3) Three features: mean, median, and max norm; 4) Four features: mean, median, max norm, and std dev.; 5) Five features: mean, median, max norm, std dev., and L1 norm; and 6) Eight features: mean, median, max norm, standard deviation, median absolute dev., mean absolute dev, L1 norm, and L2 norm.

The classification accuracy of the six cases is shown in Table VII. It is found that different feature combinations do affect the accuracy. Case 6, which is the chosen features in this paper, has better performance than other cases.

TABLE VII
INFLUENCE OF THE NUMBER OF FEATURES ON CLASSIFICATION ACCURACY FOR ALL FAULTS OF MOTOR 2 (CURRENT I_2 PROCESSED BY MP, 100% LOADING)

Machine learning methods	Sub groups	Classification accuracy in percentage using different number of features, %					
		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6 (chosen method)
SVM	Linear SVM	71.4	71.4	76.2	71	73.8	73.8
	Quadratic SVM	73.8	83	78.6	78.6	81	85.7
	Cubic SVM	92.9	90.5	90.5	90.5	88.1	88.1
	Fine Gaussian SVM	**95.2**	**97.6**	**97.6**	**97.6**	**97.6**	**97.6**
	Medium Gaussian SVM	78.6	78.6	81	81	78	81
	Coarse Gaussian SVM	73.8	71.4	66.7	66.4	66.7	71.4
KNN	**Fine KNN**	**100**	**97.6**	**97.6**	**100**	**100**	**100**
	Medium KNN	73.8	71.4	71.4	69	66.7	71.4
	Coarse KNN	16.7	16.7	16.5	16.7	16.7	16.7
	Cosine KNN	40.5	45.2	40	54.8	52.4	54.8
	Cubic KNN	73.8	71.4	71	71	71.4	71.4
	Weighted KNN	**97.6**	**100**	**100**	**100**	**100**	**100**
Ensemble	Boosted Trees	16.5	16.7	16.5	16.5	16.5	16.7
	Bagged Trees	**97.6**	**100**	**97.6**	**100**	**100**	**100**
	Subspace Discriminant	69	76.2	78	71.4	66.7	76.2
	Subspace KNN	**100**	**100**	**100**	**97.6**	**100**	**100**
	RUSBoosted Trees	16.5	16.7	16.7	16.5	16.5	16.7

VI. CALCULATED FEATURES THROUGH CURVE FITTING EQUATIONS FOR DIFFERENT MOTOR LOADINGS

In experiments, the two motors were tested under six different loadings: 100%, 85%, 70%, 50%, 30%, and 10%. However, the motor might run at a different loading under normal operation, how to obtain features for a certain loading factor when the corresponding experimental data are not available? To address this concern, curve fitting equations are developed using experimental data of the tested six loadings for a particular fault.

A. *Curve Fitting Method*

Using curve fitting, the motor loading in percentage is an independent variable; eight features processed by MP using experimental data for the six tested loadings are dependent variables. The accuracy of the developed fitting equations are evaluated by R-square values and relative errors between experimental and calculated data using these equations. The R-square value represents how closely the fitted model can follow the variance of the actual data set. It ranges from 0 to 1 where a value closer to 1 represents a better fit [49][50].

Table IX shows regression models along with their R-square values for Motor 2 with a 1BRB fault processed by MP using the stator current I_2. In these models, second order polynomial equations are adopted, x represents the percent of loading, and y represents a feature. High R-square values prove that the fitting equations follow the trend of actual measurement data. Relative errors between experimental based data and calculated data are shown in Table IX with all errors less than 8%, which further validates the accuracy of the fitting equations.

Fig. 14 shows the graphs of the eight features vs. the motor loading using the stator current I_2 for Motor 2, 1BRB fault. The dots are MP processing results using experimental data; while the solid line is determined by the curve fitting equations. Using a similar procedure, curve fitting equations for features of other types of faults can be determined.

TABLE VIII
REGRESSION MODELS FOR FEATURES USING STATOR CURRENT I_2 PROCESSED BY MP FOR MOTOR 2, 1 BRB FAULT

Feature Name	Equation	R-square Values
Mean	$y = -2E{-}07x^2 + 2E{-}05x + 0.0013$	0.9512
Median	$y = -1E{-}07x^2 + 2E{-}05x + 0.0011$	0.9197
Standard Deviation	$y = -1E{-}07x^2 + 1E{-}05x + 0.001$	0.9897
Median Absolute Value	$y = -8E{-}08x^2 + 9E{-}06x + 0.0006$	0.9168
Mean Absolute Value	$y = -8E{-}08x^2 + 1E{-}05x + 0.0008$	0.9700
L1 Norm	$y = -0.0005x^2 + 0.0549x + 3.86$	0.9512
L2 Norm	$y = -1E{-}05x^2 + 0.0012x + 0.0898$	0.9695
Maximum Norm	$y = -6E{-}07x^2 + 7E{-}05x + 0.006$	0.6482

TABLE IX
RELATIVE ERRORS BETWEEN EXPERIMENTAL BASED DATA AND CALCULATED DATA (FOR MOTOR 2, 1 BRB FAULT, STATOR CURRENT I_2)

Feature Name	Experiment based MP data	Calculated data from fitting equations	% of error
Mean (A)	0.001466	0.001480	-0.95498
Median(A)	0.001216	0.001290	-6.08553

Standard Deviation (A)	0.001130	0.001090	3.880071
Median Absolute Value (A)	0.000738	0.000682	7.588076
Mean Absolute Value (A)	0.000905	0.000892	1.425572
L1 Norm	4.399000	4.359000	0.909298
L2 Norm	0.102000	0.100800	0.689655
Maximum Norm	0.006700	0.006640	0.895522

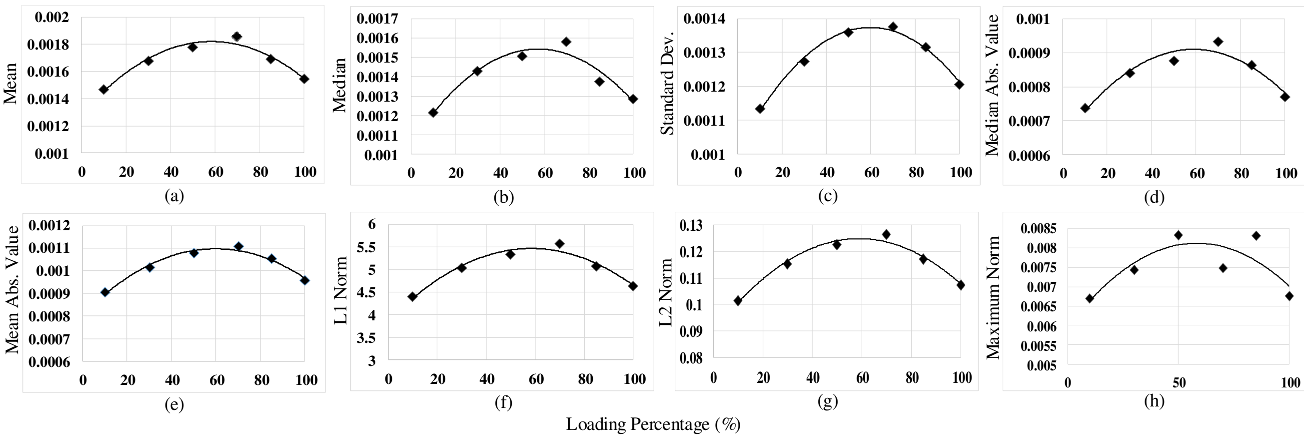


Fig. 14. Curve fitting results for features of Motor 2 with a 1BRB fault using the stator current I_2: (a) mean, (b) median, (c) standard deviation, (d) median absolute value, (e) mean absolute value, (f) L1 norm, (g) L2 norm, and (h) maximum norm.

Similarly, curve fitting can be applied to vibration signal to obtain features of a new motor loading for a fault. Table X shows the developed regression models along with their R-square values for Motor 2, 1BRB fault processed by MP using the z-axis vibration signal. In these models, the second order polynomial equations are chosen for fitting equations, x represents the percent of loading, and y represents a feature. Relative errors between experimental based data and calculated data by curve fitting equations are shown in Table XI. Fig. 15 shows the graphs of the eight features vs. the motor loading percentage in this case. The dots are MP processing results using experimental data; while the solid line is determined by the curve fitting equations.

TABLE X

REGRESSION MODELS FOR FEATURES USING Z-AXIS VIBRATION SIGNAL PROCESSED BY MP FOR MOTOR 2, 1 BRB FAULT

Feature Name	Equation	R-square Value
Mean	$y = 1E\text{-}07x^2 - 2E\text{-}05x + 0.0027$	0.9855
Median	$y = 9E\text{-}08x^2 - 1E\text{-}05x + 0.0023$	0.9898
Standard Deviation	$y = 8E\text{-}08x^2 - 1E\text{-}05x + 0.002$	0.9334
Median Absolute Value	$y = 5E\text{-}08x^2 - 8E\text{-}06x + 0.0013$	0.9615
Mean Absolute Value	$y = 6E\text{-}08x^2 - 9E\text{-}06x + 0.0016$	0.9349
L1 Norm	$y = 0.0003x^2 - 0.0495x + 8.1017$	0.9855
L2 Norm	$y = 8E\text{-}06x^2 - 0.0011x + 0.1856$	0.9707
Maximum Norm	$y = 1E\text{-}06x^2 - 0.0001x + 0.0138$	0.9345

TABLE XI

RELATIVE ERRORS BETWEEN EXPERIMENTAL BASED DATA AND CALCULATED DATA (FOR MOTOR 2, 1 BRB FAULT, Z-AXIS VIBRATION SIGNAL)

Feature Name	Simulated Value	Calculated Value	% of error
Mean (A)	0.002557	0.002510	1.840
Median(A)	0.002150	0.002209	-2.740
Standard Deviation (A)	0.001940	0.001908	1.800
Median Absolute Value (A)	0.001273	0.001225	3.770
Mean Absolute Value (A)	0.001550	0.001516	2.070
L1 Norm	7.672000	7.636700	0.460
L2 Norm	0.176000	0.175400	0.284
Maximum Norm	0.012310	0.012900	-4.790

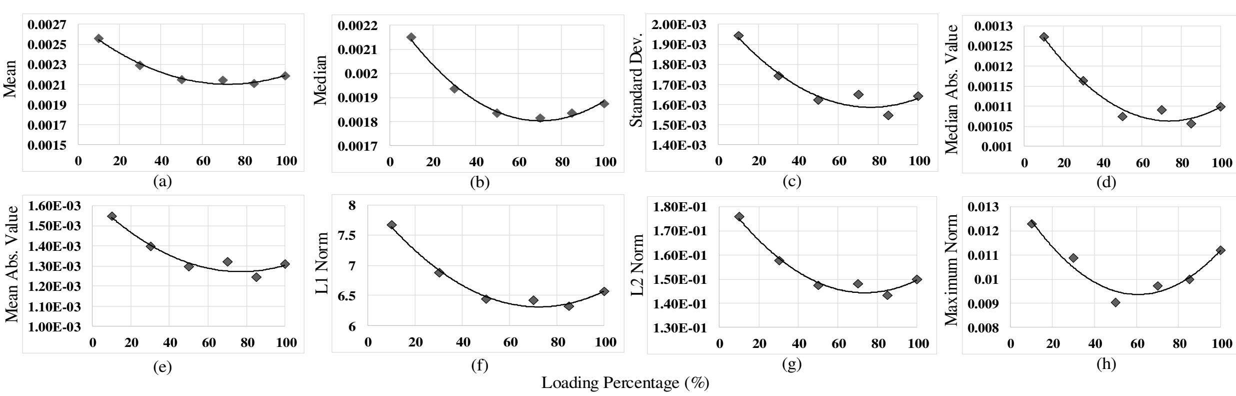


Fig. 15. Curve fitting results for features of Motor 2, 1BRB fault using the z-axis vibration signal: (a) mean, (b) median, (c) standard deviation, (d) median absolute value, (e) mean absolute value, (f) L1 norm, (g) L2 norm, and (h) maximum norm.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications

B. Machine Learning Results Using Fitting Equations

Using the developed curve fitting equations, features are calculated for three loadings (90%, 60% and 20%) that have not been tested during experiments for Motor 2. It is found that all faults can be classified at mostly 100% accuracy using the calculated features for Fine Gaussian SVM, Fine KNN, Weighted KNN, Bagged trees, and Subspace KNN. Fig. 16 shows fault classification accuracy for the three loadings for Motor 2 with the current I_2. Curve fitting equations offer effective calculation of unknown features for various motor loadings.

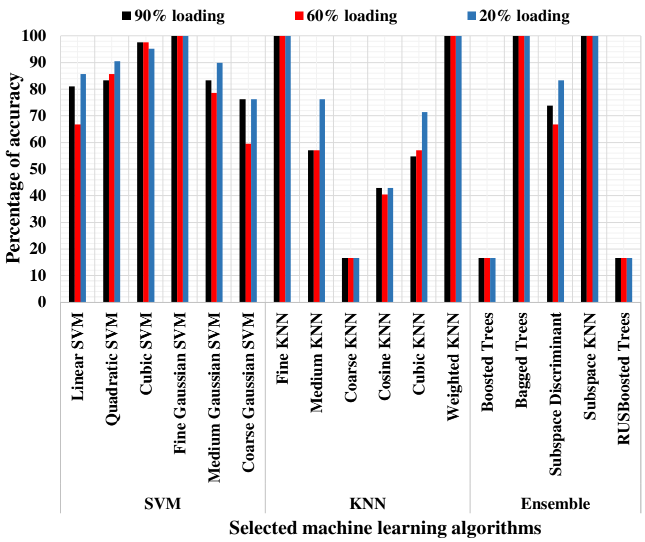


Fig. 16. Classification accuracy for all faults using features calculated by curve fitting equations for three loadings (90%, 60% and 20%) that has never been tested by experiments (Motor 2, the stator current I_2).

VII. Conclusion

Due to applications of induction motors in critical industrial processes, accurately detect various electrical or mechanical faults of induction motors are very important to avoid process down-time and large financial losses. In this paper, a machine learning based fault diagnosis method for single- and multi-faults of induction motors is proposed, developed, and validated using experimental data measured in the lab.

The following conclusions are drawn through this research: 1) The proposed fault diagnosis method is proved to be effective; 2) Either MP or DWT can be used for signal processing to extract features with a comparable accuracy; 3) The paper conducts a quantitative comparison by using stator currents and vibration signals for fault diagnosis, it is found that either stator currents or vibration signals can be used to detect the same groups of faults with a similar accuracy; 4) The number of features have influence on classification accuracy, so they should be evaluated carefully; 5) The developed curve fitting equations offer an effective calculation method of unknown features for the motors that experimental data are not available under certain loading conditions; 6) Five classifiers, Fine Gaussian SVM, fine KNN, weighted KNN, Bagged Trees, and subspace KNN, selected from MATLAB Classification Learner toolbox have mostly 100% classification accuracy for all faults of each motor, therefore, any of these five classifier can be used for induction motor fault diagnosis.

The future work for this research is to investigate how to apply the proposed fault diagnosis method to sister units of the test motor with adequate accuracy.

References

[1] X. Liang, and K. Edomwandekhoe, "Condition Monitoring Techniques for Induction Motors", *Proceedings of 2017 IEEE Industry Applications Society (IAS) Annual Meeting*, pp. 1-10, 2017.

[2] F. Filippetti, G. Franceschini, and C. Tassoni, "Neural networks aided on-line diagnostics of induction motor rotor faults", *IEEE Trans. Industry Applications*, Vol.31, No.4, pp. 892 – 899, 1995.

[3] R. R. Schoen, B. K. Lin, T. G. Habetler, J. H. Schlag, and S. Farag, "An unsupervised, on-line system for induction motor fault detection using stator current monitoring", *IEEE Trans. Industry Applications*, Vol. 31, No. 6, pp. 1280 - 1286, 1995.

[4] J. F. Martins, V. Ferno Pires, and A. J. Pires, "Unsupervised neural-network-based algorithm for an on-line diagnosis of three-phase induction motor stator fault", *IEEE Trans. Industrial Electronics,* Vol. 54, No. 1, pp. 259 – 264, 2007.

[5] H. Nejjari, and M. E. H. Benbouzid, "Monitoring and diagnosis of induction motors electrical faults using a current Park's vector pattern learning approach", *IEEE Trans. Industry Applications*, Vol. 36, No. 3, pp. 730 - 735, 2000.

[6] S. Wu, and T. W. S. Chow, "Induction machine fault detection using SOM-based RBF neural networks", *IEEE Trans. Industrial Electronics,* Vol. 51, No. 1, pp. 183 - 194, 2004.

[7] H. Su, and K. T. Chong, "Induction machine condition monitoring using neural network modeling", *IEEE Trans. Industrial Electronics,* Vol. 54, No. 1, pp. 241 - 249, 2007.

[8] T. Boukra, A. Lebaroud, and G. Clerc, "Statistical and neural-network approaches for the classification of induction machine faults using the ambiguity plane representation", *IEEE Trans. Industrial Electronics,* Vol. 60, No. 9, pp. 4034 - 4042, 2013.

[9] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, "Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks", *IEEE Trans. Industrial Electronics,* Vol. 60, No. 8, pp. 3398 - 3407, 2013.

[10] W. Sun, R. Zhao, R. Yan, S. Shao, and X. Chen, "Convolutional discriminative feature learning for induction motor fault diagnosis", *IEEE Trans. Industrial Informatics*, Vol. 13, No. 3, pp. 1350 – 1359, 2017.

[11] P. V. Goode, and M. Chow, "Using a neural/fuzzy system to extract heuristic knowledge of incipient faults in induction motors. Part I-methodology", *IEEE Trans. Industrial Electronics*, Vol. 42, No. 2, pp. 131 - 138, 1995.

[12] P. V. Goode, and M. Chow, "Using a neural/fuzzy system to extract heuristic knowledge of incipient faults in induction motors. Part II-application", *IEEE Trans. Industrial Electronics*, Vol. 42, No. 2, pp. 139 - 146, 1995.

[13] S. Altug, M. Chen, and H. J. Trussell, "Fuzzy inference systems implemented on neural architectures for motor fault detection and diagnosis", *IEEE Trans. Industrial Electronics*, Vol. 46, No. 6, pp. 1069 - 1079, 1999.

[14] M. S. Ballal, Z. J. Khan, H. M. Suryawanshi, and R. L. Sonolikar, "Adaptive neural fuzzy inference system for the detection of inter-turn insulation and bearing wear faults in induction motor", *IEEE Trans. Industrial Electronics,* Vol. 54, No. 1, pp. 250 - 258, 2007.

[15] M. Seera, C. P. Lim, D. Ishak, and H. Singh, "Fault detection and diagnosis of induction motors using motor current signature analysis and a hybrid FMM–CART Model", *IEEE Trans. Neural Networks and Learning Systems*, Vol. 23, No.1, pp. 97 – 108, 2012.

[16] P. J. C. Branco, J. A. Dente, and R. V. Mendes, “Using immunology principles for fault detection”, *IEEE Trans. Industrial Electronics,* Vol. 50, No. 2, pp. 362 - 373, 2003.

[17] O. Ondel, E. Boutleux, E. Blanco, and G. Clerc, “Coupling pattern recognition with state estimation using Kalman filter for fault diagnosis”, *IEEE Trans. Industrial Electronics,* Vol. 59, No. 11, pp. 4293 - 4300, 2012.

[18] R. Razavi-Far, M. Farajzadeh-Zanjani, and M. Saif, “An integrated class-imbalanced learning scheme for diagnosing bearing defects in induction motors”, *IEEE Trans. Industrial Informatics*, Vol. 13, No. 6, pp. 2758 - 2769, 2017.

[19] C. Sun, M. Ma, Z. Zhao, and X. Chen, “Sparse deep stacking network for fault diagnosis of motor”, *IEEE Trans. Industrial Informatics*, Vol. 14, No. 7, pp. 3261 - 3270, 2018.

[20] F. B. Abid, S. Zgarni, and A. Braham, “Distinct bearing faults detection in induction motor by a hybrid optimized SWPT and aiNet-DAG SVM”, *IEEE Trans. Energy Conversion*, DOI 10.1109/TEC.2018.2839083 (early access).

[21] I. Martin-Diaz, D. Morinigo-Sotelo, O. Duque-Perez, and R. J. Romero-Troncoso, “An experimental comparative evaluation of machine learning techniques for motor fault diagnosis under various operating conditions”, *IEEE Trans. Industry Applications*, Vol. 54, No. 3, pp. 2215 – 2224, 2018.

[22] W. F. Godoy, I. N. d. Silva, A. Goedtel, R. H. C. Palácios, and T. D. Lopes, “Application of intelligent tools to detect and classify broken rotor bars in three-phase induction motors fed by an inverter”, *IET Electric Power Applications*, Vol. 10, No. 5, pp. 430 – 439, 2016.

[23] J. Seshadrinath, B. Singh, and B. K. Panigrahi, “Investigation of vibration signatures for multiple fault diagnosis in variable frequency drives using complex wavelets”, *IEEE Trans. Power Electronics*, Vol. 29, No. 2, pp. 936 – 945, 2014.

[24] M. Pineda-Sanchez, M. Riera-Guasp, J. A. Antonino-Daviu, J. Roger-Folch, J. Perez-Cruz, and R. Puche-Panadero, “Diagnosis of induction motor faults in the fractional Fourier domain”, *IEEE Trans. Instrumentation and Measurement*, Vol. 59, No. 8, pp. 2065 – 2075, 2010.

[25] O. A. Mohammed, N. Y. Abed, and S. C. Ganu, “Modeling and characterization of induction motor internal faults using finite-element and discrete wavelet transforms”, *IEEE Trans. Magnetics*, Vol. 42, No. 10, pp. 3434 – 3436, 2006.

[26] A. Bouzida, O. Touhami, R. Ibtiouen, A. Belouchrani, M. Fadel, and A. Rezzoug, “Fault Diagnosis in Industrial Induction Machines Through Discrete Wavelet Transform”, *IEEE Trans. Industrial Electronics,* Vol. 58, No. 9, pp. 4385 - 4395, 2011.

[27] S. G. Mallat, and Z. Zhang, “Matching pursuits with time-frequency dictionaries”, IEEE Trans. Signal Processing, Vol. 41, No. 12, pp. 3397 – 3415, 1993.

[28] J. Li, M. Li, X. Yao, and H. Wang, “An Adaptive Randomized Orthogonal Matching Pursuit Algorithm with Sliding Window for Rolling Bearing Fault Diagnosis,” *IEEE Access*, pp. 1–1, 2018.

[29] S. C. Ks, A. Mishra, V. Shirhatti, and S. Ray, “Comparison of Matching Pursuit Algorithm with Other Signal Processing Techniques for Computation of the Time-Frequency Power Spectrum of Brain Signals,” *Journal of Neuroscience*, vol. 36, no. 12, pp. 3399–3408, 2016.

[30] K. Edomwandekhoe and X. Liang, “Advanced feature selection for broken rotor bar faults in induction motors,” *2018 IEEE/IAS 54th Industrial and Commercial Power Systems Technical Conference (I&CPS)*, 2018.

[31] J. Seshadrinath, B. Singh, and B. K. Panigrahi, “Incipient broken rotor bar detection in induction motors using vibration signals and the orthogonal matching pursuit algorithm”, *IEEE Trans. Power Electronics*, Vol. 29, No. 2, pp. 936 – 945, 2014.

[32] “wmpdictionary - Dictionary for matching pursuit”, https://www.mathworks.com/help/wavelet/ref/wmpdictionary.html. Visited on July 22, 2018 at 3:49 pm.

[33] J. J. Saucedo-Dorantes, M. Delgado-Prieto, R. A. Osornio-Rios, and R. D. J. Romero-Troncoso, “Multifault Diagnosis Method Applied to an Electric Machine Based on High-Dimensional Feature Reduction,” *IEEE Trans. Industry Applications*, vol. 53, no. 3, pp. 3086–3097, 2017.

[34] M. F. Al-Saleh and A. E. Yousif, “Properties of the Standard Deviation that are Rarely Mentioned in Classrooms,” *Austrian Journal of Statistics*, vol. 38, no. 3, p. 193, Mar. 2016.

[35] M. Aktas and V. Turkmenoglu, “Wavelet-based switching faults detection in direct torque control induction motor drives,” *IET Science, Measurement & Technology*, vol. 4, no. 6, pp. 303–310, Jan. 2010.

[36] M. Riera-Guasp, J. Antonino-Daviu, M. Pineda-Sanchez, R. Puche-Panadero, and J. Perez-Cruz, “A General Approach for the Transient Detection of Slip-Dependent Fault Components Based on the Discrete Wavelet Transform,” *IEEE Trans. Industrial Electronics*, vol. 55, no. 12, pp. 4167–4180, 2008.

[37] J. Antonino-Daviu, M. Riera-Guasp, J. Roger-Folch, and M. P. Molina, “Validation of a new method for the diagnosis of rotor bar failures via wavelet transformation in industrial induction machines,” *IEEE Trans. Ind. Appl.*, vol. 42, no. 4, pp. 990–996, 2006.

[38] S. S., K. U. Rao, R. Umesh, and H. K. S., “Condition monitoring of Induction Motor using statistical processing,” *2016 IEEE Region 10 Conference (TENCON)*, 2016.

[39] P. Gomez-Gil, J. Rangel-Magdaleno, J. M. Ramirez-Cortes, E. Garcia-Trevino, and I. Cruz-Vega, “Intelligent identification of induction motor conditions at several mechanical loads,” *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, 2016.

[40] K.-R. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, “An introduction to kernel-based learning algorithms,” *IEEE Trans. Neural Networks*, vol. 12, no. 2, pp. 181–201, 2001.

[41] Xiaodong Liang, Yi He, Massimo Mitolo, and Weixing Li, "Support Vector Machine Based Dynamic Load Model Using Synchrophasor Data", Proceedings of IEEE 54th Industrial and Commercial Power Systems (I&CPS) Conference, pp. 1-11, May 2018.

[42] H. Jung, S.-W. Kang, M. Song, S. Im, J. Kim, and C.-S. Hwang, “Towards Real-Time Processing of Monitoring Continuous k-Nearest Neighbor Queries,” *Frontiers of High Performance Computing and Networking – ISPA 2006 Workshops Lecture Notes in Computer Science*, pp. 11–20, 2006.

[43] M. Wiering and H. V. Hasselt, “Ensemble Algorithms in Reinforcement Learning,” *IEEE Trans. Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 4, pp. 930–936, 2008.

[44] L. Breiman, “Bagging predictors,” *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[45] C. K. Tham, “Reinforcement learning of multiple tasks using a hierarchical CMAC architecture,” *Robotics and Autonomous Systems*, vol. 15, no. 4, pp. 247–274, 1995.

[46] R. Sun and T. Peterson, “Multi-agent reinforcement learning: weighting and partitioning,” *Neural Networks*, vol. 12, no. 4-5, pp. 727–753, 1999.

[47] T. Wong and N. Yang, "Dependency Analysis of Accuracy Estimates in k-Fold Cross Validation," in *IEEE Trans. Knowledge and Data Engineering*, vol. 29, no. 11, pp. 2417-2427, 1 Nov. 2017.

[48] I. Martin-Diaz, D. Morinigo-Sotelo, O. Duque-Perez and R. D. J. Romero-Troncoso, "Advances in Classifier Evaluation: Novel Insights for an Electric Data-Driven Motor Diagnosis," in *IEEE Access*, vol. 4, pp. 7028-7038, 2016.

[49] C. Andalib-Bin-Karim, X. Liang, N. Khan, and H. Zhang, “Determine Q–V Characteristics of Grid-Connected Wind Farms for Voltage Control Using a Data-Driven Analytics Approach,” *IEEE Trans. Industry Applications*, vol. 53, no. 5, pp. 4162–4175, 2017.

[50] M. Z. Ali, M. N. S. K. Shabbir, X. Liang, Y. Zhang, and T. Hu, "Experimental Investigation of Machine Learning Based Fault Diagnosis for Induction Motors," *Proceedings of 2018 IEEE Industry Applications Society (IAS) Annual Meeting*, pp. 1-14, Portland, OR, USA, September 23 - 27, 2018.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIA.2019.2895797, IEEE Transactions on Industry Applications

Mohammad Zawad Ali (S'17) was born in Khulna, Bangladesh. He received his Bachelor of Science degree in Electrical and Electronic Engineering from Islamic University of Technology, Gazipur, Bangladesh, in 2013. His research interests include power systems, renewable energy, conditioning monitoring and fault diagnosis of induction motors.

During his second year of undergrad program, he worked as a trainee in Khulna Power Station of Bangladesh Power Development Board, Khulna, Bangladesh from November 1 to 30, 2011. From May 2014 to August 2017, he served as a Lecturer at Dhaka International University, Dhaka, Bangladesh. Since September 2017, he has enrolled in the Master of Engineering program at Memorial University of Newfoundland, St. John's, NL, Canada, where he is currently pursuing his M.Eng degree.

Md Nasmus Sakib Khan Shabbir (S'17) was born in Tangail, Bangladesh. He received his B.Sc. degree in electrical and electronic engineering from Rajshahi University of Engineering and Technology, Bangladesh, in 2012.

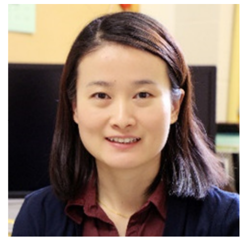
From 2013 to 2016, he served as a Lecturer in the department of Electrical and Electronic Engineering at Bangladesh University of Business and Technology, and was promoted to the rank of Assistant Professor in 2016 in the aforesaid department and served until October in 2017. He is currently pursuing his PhD degree at Memorial University of Newfoundland, St. John`s, Canada. His research interest includes the design and control of high performance power electronic converters and their applications in renewable energy integration to power grids.

Xiaodong Liang (M'06–SM'09) was born in Lingyuan, China. She received her B.Eng. and M.Eng. degrees from Shenyang Polytechnic University, Shenyang, China in 1992 and 1995, respectively, the M.Sc. degree from the University of Saskatchewan, Saskatoon, Canada in 2004, and the Ph.D. degree from the University of Alberta, Edmonton, Canada in 2013, all in Electrical Engineering. Her research interests include power system dynamics, renewable energy, and electric machines.

From 1995 to 1999, she served as a lecturer at Northeastern University, Shenyang, China. In October 2001, she joined Schlumberger in Edmonton, Canada, and was promoted to be a Principal Power Systems Engineer with this world's leading oil field service company in 2009. After serving Schlumberger for almost 12 years, she joined Washington State University in Vancouver, Washington, United States in August 2013. From August 2013 to May 2015, she was an Assistant Professor at Washington State University. In July 2015, she joined Memorial University of Newfoundland in St. John's, Canada, where she is currently an Associate Professor.

Dr. Liang is a registered professional engineer in the province of Newfoundland and Labrador, Canada.

Yu Zhang was born in Zhengzhou, Henan, China. She received the B.Eng degree in Computer Science from Hunan University, Changsha, China in 2014, and the M.Eng degree in Computer Science from Beijing University of Posts and Telecommunications, Beijing, China in 2017. Currently, she is pursuing her M.Sc degree in Computer Science at Memorial University of Newfoundland, St. John's, Canada. Her research interests include evolutionary computing, machine learning, complex network, and bioinformatics.



Ting Hu received her B.Sc. in Applied Mathematics and M.Sc. in Computer Science from Wuhan University, China in 2003 and 2005, respectively, and her Ph.D. in Computer Science from Memorial University, Canada in 2010. Her postdoctoral training in Computational Genetics was conducted at the Geisel School of Medicine at Dartmouth College, U.S.A. Her research interests include areas of computational intelligence and bioinformatics.

Dr. Hu joined Memorial University as an Assistant Professor at the Department of Computer Science in 2015. She has served the editorial board of international journals including Genetic Programming and Evolvable Machines, Entropy, Genes, and has chaired international conferences including European Conference on Genetic Programming (EuroGP), Genetic and Evolutionary Computation Conference (GECCO) GP track, European Conference on Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics (EvoBio).