INDUCTION MOTORS FAULT DIAGNOSIS USING MACHINE LEARNING AND ADVANCED SIGNAL PROCESSING TECHNIQUES

By

Mohammad Zawad Ali

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Engineering

Faculty of Engineering and Applied Science

Memorial University of Newfoundland

October 2019

St. John's   Newfoundland and Labrador   Canada

Abstract

In this thesis, induction motors fault diagnosis are investigated using machine learning and advanced signal processing techniques considering two scenarios: 1) induction motors are directly connected online; and 2) induction motors are fed by variable frequency drives (VFDs). The research is based on experimental data obtained in the lab. Various single- and multi- electrical and/or mechanical faults were applied to two identical induction motors in experiments. Stator currents and vibration signals of the two motors were measured simultaneously during experiments and were used in developing the fault diagnosis method. Signal processing techniques such as Matching Pursuit (MP) and Discrete Wavelet Transform (DWT) are chosen for feature extraction. Classification algorithms, including decision trees, support vector machine (SVM), K-nearest neighbors (KNN), and Ensemble algorithms are used in the study to evaluate the performance and suitability of different classifiers for induction motor fault diagnosis. Novel curve or surface fitting techniques are implemented to obtain features for conditions that have not been tested in experiments. The proposed fault diagnosis method can accurately detect single- or multi- electrical and mechanical faults in induction motors either directly online or fed by VFDs.

In addition to the machine learning method, a threshold method using the stator current signal processed by DWT is also proposed in the thesis.

Acknowledgements

Table of Contents

List of Tables

List of Figures

List of Abbreviations

IEEE	Institute of Electrical and Electronics Engineers
EPRI	Electric Power Research Institute
IM	Induction Motor
MP	Matching Pursuit
DWT	Discrete Wavelet Transform
AI	Artificial Intelligence
SVM	Support Vector Machine
KNN	K-Nearest Neighbors
MSCA	Motor Current Signature Analysis
ANN	Artificial Neural Network
Op-SWPT	Optimized Stationary Wavelet Packet Transform
AIS	Artificial Immune System
VFD	Variable Frequency Drive
UNB	Unbalance (shaft rotation)
UV	Unbalance Voltage
BF	Bearing Fault
BRB	Broken Rotor Bar
OMP	Orthogonal Matching Pursuit
RBF	Radial Basis Function
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
TPR	True Positive Rate
FPR	False Positive Rate
UT	Universal Threshold

List of symbols

x_i	The ith sampled measurement point, $i = 1, 2, 3,\ldots,$ N
μ_X	Mean
med	Median
Median_AD	Median Absolute Deviation
Mean_AD	Mean Absolute Deviation
σ	Standard Deviation (Std. Dev.)
N_{Ls}	Total number of decomposition level
f_s	Sampling frequency
f_{dj}	Frequency level at detail coefficient
f_{aj}	Frequency level at approximations coefficient
D	The set of first level detail coefficients
E_j	Energy of each frequency band
N	Number of samples

Note: Other symbols not mentioned are defined in the text.

Chapter 1

Introduction

1.1 Background Study

Induction motors are workhorse for various industrial sectors. General applications of induction motors include pumps, conveyors, machine tools, centrifugal machines, presses, elevators, and packaging equipment. They are also used in hazardous locations such as petrochemical and natural gas plants, and in severe environment such as grain elevators, shredders, and equipment for coal plants [1]-[3].

Although relative robust, induction motors are still susceptible to many types of faults. A motor failure that is not identified in an initial stage may become catastrophic, the motor may suffer severe damage and it may cause production shutdowns. Such shutdowns are costly in terms of lost production time, maintenance costs, and wasted raw materials. For reliable and smooth operation in any industrial process, it is important to know the distribution of different failure sections of an induction motor for condition monitoring and incipient fault diagnosis.

Recently, research has designated different fault distributions of induction motors within a range of 0.75 kW to 150 kW and provided probable scenario and decision processes to diagnose those faults. The main distribution of induction motor faults, shown in Fig. 1.1, can be categorized as rotor bar (7%), stator winding (21%), bearing fault (69%), and shaft/coupling and others (3%). Approximately two-third of faults arose due to bearing and one-fifth occurred due to stator windings [1] [2].

A similar statistical survey conducted by IEEE and electric power research institute (EPRI) on induction motor faults and the percentage of different faults with respect to the total number of faults is tabulated in Table 1.1 [4].
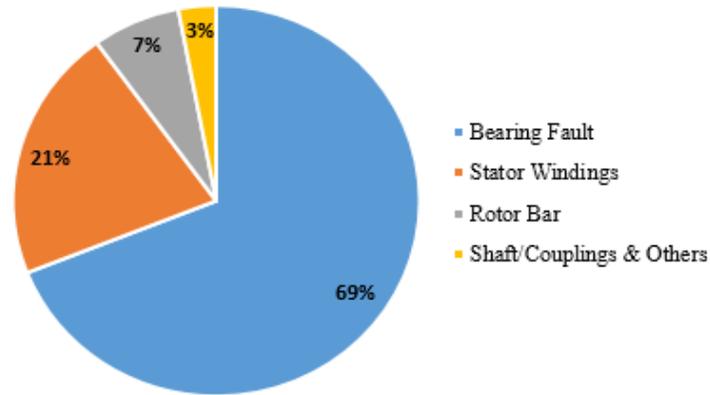
Fig. 1. 1. Different fault distribution of an induction motor [2].

The maintenance of electrical machines can be categorized in three ways: 1) condition-based maintenance, 2) scheduled maintenance, and 3) breakdown maintenance. Condition-based maintenance includes noticing and receiving several periodic data like voltage, current spectrum, torque profile etc. from a machine during operation and taking necessary steps to prevent any fault at the initial stage to minimize the machine's downtime [4]. The scheduled maintenance is defined as employing expertise to stop the machine for checkup, determine any defects and repair and replace the part accordingly, although the process requires a long downtime. Lastly, the breakdown maintenance, which occurs during a mechanical failure of the machine, requires the replacement of the whole machine rather than fixing or replacing the faulty parts of the machine [3]-[5]. Therefore, it is advantageous to implement a condition-based monitoring system because it requires less maintenance, lowers the cost and reduces the downtime of the machine [5].

Table 1. 1: Statistical survey results for induction motor faults by IEEE and EPRI [3] [4].

Major fault components	Studied by	
	IEEE (%)	EPRI (%)
Rotor fault	8	9
Bearing fault	42	41
Stator fault	28	36
Others	22	14

The motor faults are due to mechanical and electrical stresses. Mechanical stresses are caused by overloads and abrupt load changes, which can produce bearing faults and rotor bar breakage. Electrical stresses are usually associated with the power supply. Induction motors can be energized from constant frequency sinusoidal power supplies or from adjustable speed ac drives.

However, induction motors are more susceptible to faults when supplied by ac drives. This is due to the extra voltage stress on the stator windings, the high frequency stator current components, and the induced bearing currents, caused by ac drives. In addition, motor over-voltages can occur because of the length of cable connections between a motor and an ac drive. Such electrical stresses may produce stator winding short circuits and result in a complete motor failure.

The aim of the condition monitoring process of an induction motor is to demonstrate a reliable mechanism for fault detection at the initial stage so that necessary steps can be taken [6]. According to the different survey processes, different types of fault are revealed under different conditions that are shown in Fig. 1.2 [3], [5]-[7]. Observing and examining the corresponding abnormalities in induction motor voltage, current spectrum and leakage flux helps to monitor the condition and diagnose the fault at the incipient breakdown stage of an induction motor [7].

There are several methods used by researchers to investigate the diagnosis process of faulty machines, including their bearing faults, broken rotor bars, air gap eccentricity and stator winding inter turn faults [5]. The objective of this thesis is to obtain an effective fault diagnosis method using machine learning and advanced signal processing techniques.
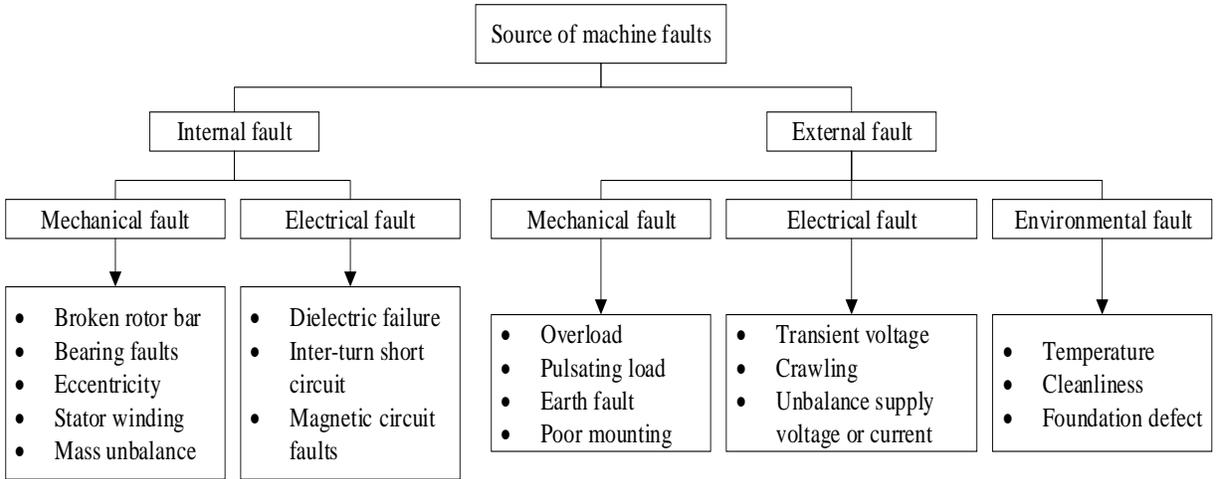
Fig. 1. 2. Summary of different faults under different operating conditions [3], [5]-[7].

1.2 Thesis Outline

This thesis consists of three manuscripts, two have been already published, and one has been accepted for publication.

Chapter 1

In Chapter 1, the importance of the research topic is introduced; the objectives of the research work are described.

Chapter 2

In Chapter 2, a literature review for the research work is conducted. The main objective of this thesis is to develop a fault diagnosis method for induction motors using machine learning and advanced signal processing techniques.

Chapter 3

In Chapter 3, a machine learning based fault diagnosis method is proposed for induction motors using experimental data for induction motors connected directly online. Various single- and multi- electrical and/or mechanical faults were applied to two identical induction motors in experiments. Stator currents and

vibration signals of the motors were 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 the MATLAB Classification Learner toolbox, are used in the study to evaluate the performance and suitability of different classifiers for induction motor fault diagnosis. A novel curve fitting technique is developed to calculate features for the conditions that are not tested in experiments. The proposed fault diagnosis method can accurately detect single- or multi- electrical and mechanical faults in induction motors. A version of this chapter has been published in the IEEE Transactions on Industry Applications in May/June 2019 regular issue.

Chapter 4

In Chapter 4, a general methodology is developed by using experimentally measured stator current signals under the full load condition of an induction motor connected directly online. The measured stator current for various single- and multi-electrical faults of the induction motor are investigated to obtain signatures for fault diagnosis. In this chapter, the DWT is chosen for signal processing. The threshold and energy values at each decomposition level for the DWT analysis are evaluated. A version of this chapter has been published in the proceedings of 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE).

Chapter 5

In Chapter 5, a machine learning based fault diagnosis method is developed for induction motors fed by VFDs. Two identical 0.25 HP induction motors under healthy, single- and multi-fault conditions were tested in the lab with different VFD output frequencies settings and motor loadings. The stator current and vibration signals of the motor were recorded simultaneously under steady-state for each test. Both signals have been evaluated for their suitability for machine learning. The DWT is chosen to process the signals

and extract the features. Four families of algorithms from the MATLAB Classification Learner Toolbox, decision trees, SVM, KNN, and Ensemble, with twenty classifiers, are used to evaluate their classification accuracy. To allow fault diagnosis for untested motor operating conditions with different combinations of the motor operating frequency and the motor loading factor, feature calculation formulas are developed through surface fitting for the conditions that are not tested in experiments. The proposed fault diagnosis method can accurately detect single- or multi- electrical and mechanical faults in induction motors fed by VFD. A version of this chapter has been accepted by 2019 IEEE Industry Application Society (IAS) Annual Meeting.

Chapter 6

In Chapter 6, the research outcomes are summarized and the potential future research scope is addressed.

References

[1] A. H. Bonnett and C. Yung, "Increased efficiency versus increased reliability," *IEEE Industry Applications Magazine*, vol. 14, 2008.

[2] X. Wen, "A hybrid intelligent technique for induction motor condition monitoring," University of Portsmouth, 2011.

[3] S. Karmakar, S. Chattopadhyay, M. Mitra, and S. Sengupta, "Induction Motor Fault Diagnosis," *Publisher Springer Singapore*, 2016.

[4] P. Albrecht, J. Appiarius, E. Cornell, D. Houghtaling, R. McCoy, E. Owen*, et al.*, "Assessment of the reliability of motors in utility applications," *IEEE transactions on energy conversion,* pp. 396-406, 1987.

[5] X. Liang and K. Edomwandekhoe, "Condition monitoring techniques for induction motors," *2017 IEEE Industry Applications Society Annual Meeting*, 2017.

[6] S. A. Taher and M. Malekpour, "A novel technique for rotor bar failure detection in single-cage induction motor using FEM and MATLAB/SIMULINK," *Mathematical Problems in Engineering,* vol. 2011, 2011.

[7] K. M. B. Gamal, S. Handa, and M. R. Murthy, "A Fault Diagnostic Method for Monitoring Induction Motor," *International Journal of Engineering Science,* vol. 4610, 2017.

Chapter 2

Literature Review

2.1 Streams of Research on Fault Diagnosis of the Induction Motor

Induction motors play a vital role in safe and efficient running of industrial plants and processes due to their low cost, strength and economical maintenance. Early detection of abnormalities in induction motors will help to avoid destructive failures. The goal of machine condition monitoring and fault diagnosis is to provide useful and effective information on the condition of equipment in a timely manner [1]-[4].

There are many published techniques and commercially available tools to monitor induction motors to assure a high degree of reliable uptime. Redundancy and conservative design systems have been adopted for improving the reliability of induction motor drive systems for a variety of faults that can occur, but these techniques are expensive to implement.

There are several advantages of condition monitoring of induction motors can be addressed as follows: (i) improved operating efficiency, (ii) reduced maintenance costs for better planning, (iii) extended operational life of the machine, and (iv) increased machine availability and reliability etc. However, there are some disadvantages that must be weighed in the decision to use machine condition monitoring and fault diagnostics and these drawbacks includes: (i) monitoring equipment costs, (ii) a significant running time to collect machine histories and trends is usually needed, and (iii) operational costs etc. [1][2]. Condition monitoring leading to incipient fault detection and prediction for induction motors has attracted many researchers in the past few years owing to its considerable influence on the safe operation of many industrial processes. It is important to be able to detect motor faults while they are still developing [2]. The one-line diagram of a general approach to condition monitoring for the induction motor is shown in Fig. 2.1 [3].
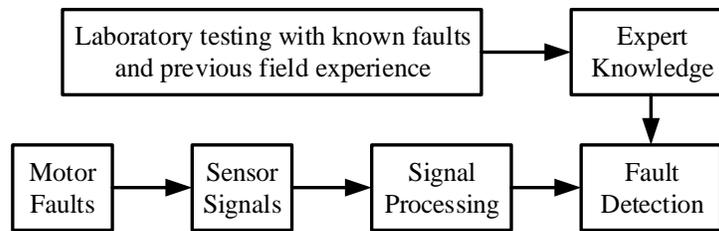
Fig. 2. 1. General approach of condition monitoring [3].

Early detection and precise diagnosis of incipient faults allow preventive maintenance to be performed and provide sufficient time for controlled shutdown of the affected process. They can reduce financial losses and avoid catastrophic consequences. An ideal condition monitoring and fault diagnosis scheme should take the minimum measurements from a motor. A condition can be inferred to give a clear indication of incipient fault modes in a minimal time. Condition monitoring establishes a map between input signals and output indications of motor conditions. Classifying the type of motor faults and determining the severity of faults are not easy because they are affected by many factors [4].

In the literature, there are three streams of research on fault diagnosis for induction motors, which can be categorized as follows: 1) signature extraction based approaches, 2) model-based approaches, and 3) knowledge-based approaches. The signature extraction based approaches quantify fault signatures in time and/or the 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 of 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 ring faults. However, the false fault indication is a common issue with MSCA that needs to be improved. The model-based approaches depend on mathematical modeling to predict behaviors of induction motors under fault conditions. Although model-based approaches can provide warnings and estimate incipient faults, their 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. The knowledge-based approaches use machine learning techniques for on-line and off-line applications. Artificial intelligence (AI) 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 is emerging as a promising research direction for induction motor fault diagnosis with great industrial application potential [2][5][6].

Signal processing is a key step for condition monitoring and fault diagnosis. It can be categorized as follows: 1) time domain analysis, 2) frequency domain analysis, and 3) time-frequency domain analysis. There are many signals, including vibration and electrical signals, for motor condition monitoring and fault diagnosis. However, an important factor for motor condition monitoring and fault diagnosis is the ability to extract the features of motor signals. The goal is to extract features which are related to specific motor fault modes. A feature extraction technique is needed for signal processing of recorded time-series signals over a long period of time to obtain suitable feature parameters for condition monitoring and fault diagnosis. By employing appropriate signal analysis algorithms, it is feasible to detect changes in signals caused by faulty components. The aim of feature extraction is to extract the signal features hidden in the original time domain. Corresponding to different signals, a signal analysis method should be properly selected such that the feature value of signals can be boosted to improve diagnostic sensitivity to a motor fault.

Most of the analysis used for fault diagnosis, starting about three decades ago, was performed using fast Fourier transform (FFT) based tools on the motor current or vibration signature. However, FFT has some limitations, like the masking of characteristic frequencies (generally small frequency) by supply frequency, inappropriateness for transient signals, etc. To overcome these limitations, different new techniques are being used now [1]-[3]. Some of the present signal processing techniques are reported in the literature as critical steps for fault diagnosis. These techniques include short-time Fourier transform (STFT), Wigner-

Ville distribution (WVD), power spectral density (PSD), wavelet transform (WT) [6]-[10], multiple signal classification (MUSIC) method [11]-[13], Hilbert transform [14]-[16], or hybrid techniques, such as combining Wavelet and Hilbert transforms with a linear discrimination method [17], and homogeneity analysis with the Gaussian probability density function [18].

The condition monitoring and fault diagnosis of induction motors have moved from traditional techniques to AI techniques. The knowledge-based approach using AI and machine learning opens a pathway to exciting new research directions in condition monitoring and fault diagnosis of induction motors. During the 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 [19]-[32]. As one appealing feature of ANN that can be used for on-line applications, many of the proposed ANN methods are for on-line fault diagnosis of induction motors [19]-[22]. The hybrid ANN methods include: Park's vector–neural networks approach [22], an analytical redundancy method based neural network modeling [24], statistical and neural network approaches [25][26], and the convolutional discriminative feature learning method [27]. One of the most popular hybrid ANN methods is combining ANN with fuzzy logic, which can provide accurate fault detection with heuristic interpretation [28]-[32]. These techniques use association, reasoning and decision making processes as would the human brain in solving diagnostic problems [4]. In this chapter, the literature review for the research work is focused on condition monitoring and fault diagnosis of the induction motor. The main objective of this thesis is to develop a fault diagnosis method for induction motors using machine learning and advanced signal processing techniques.

2.2 Outcomes of the Thesis

In this thesis, the fault diagnosis of three-phase squirrel-cage induction motors is investigated by processing the measured stator current and vibration signals for two sample machines in the lab through

advanced signal processing techniques and machine learning. The proposed approaches and outcomes of the thesis are described as follows:

- In chapter 3, a machine learning based fault diagnosis method is proposed for induction motors fed directly online using experimental data. MP and DWT are used for signal processing, and their performance is compared. A novel curve fitting technique is developed to calculate features for the conditions that are not tested. The proposed fault diagnosis method can accurately detect single- or multi- electrical and mechanical faults in induction motors fed directly online.

- In chapter 4, a general methodology based on threshold value is developed by DWT processing of the measured stator current signals under the full load condition of an induction motor fed directly online. The suitability of the DWT method is assessed by the threshold value of each decomposition level and the energy of each detail level. A robust fault diagnosis method is proposed for classifying various faults of induction motors based on the DWT processing results.

- In chapter 5, a machine learning based fault diagnosis method is developed for induction motors fed by VFDs using experimental data considering both motor loading and VFD output frequency. The surface fitting technique is used to determine feature calculation formula for the conditions that are not tested during experiments.

References

[1] S. Karmakar, S. Chattopadhyay, M. Mitra, and S. Sengupta, "Induction Motor Fault Diagnosis," *Publisher Springer Singapore*, 2016.

[2] X. Wen, "A hybrid intelligent technique for induction motor condition monitoring," University of Portsmouth, 2011.

[3] W. Wang and D. D. Li, "Health Condition Monitoring of Induction Motors," *Induction Motors - Applications, Control and Fault Diagnostics*, 2015.

[4] P. Tavner, "Review of condition monitoring of rotating electrical machines," *IET Electric Power Applications*, vol. 2, no. 4, p. 215, 2008.

[5] X. Liang and K. Edomwandekhoe, "Condition monitoring techniques for induction motors," *2017 IEEE Industry Applications Society Annual Meeting*, 2017.

[6] M. Z. Ali, M. N. S. K. Shabbir, X. Liang, Y. Zhang, and T. Hu, "Machine Learning based Fault Diagnosis for Single- and Multi-Faults in Induction Motors Using Measured Stator Currents and Vibration Signals," *IEEE Transactions on Industry Applications*, pp. 1–1, 2019.

[7] N. Bessous, S. E. Zouzou, S. Sbaa, and W. Bentrah, "A comparative study between the MCSA, DWT and the vibration analysis methods to diagnose the dynamic eccentricity fault in induction motors," *2017 6th International Conference on Systems and Control (ICSC)*, 2017.

[8] A. Mohanty and C. Kar, "Fault Detection in a Multistage Gearbox by Demodulation of Motor Current Waveform," *IEEE Trans. Ind. Electron.*, vol. 53, no. 4, pp. 1285–1297, 2006.

[9] 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. Ind. Electron., vol. 55, no. 12, pp. 4167–4180, 2008.

[10] A. Glowacz, "Diagnostics of DC and Induction Motors Based on the Analysis of Acoustic Signals," *Measurement Science Review*, vol. 14, no. 5, pp. 257–262, Jan. 2014.

[11] R. Romero-Troncoso, A. Garcia-Perez, D. Morinigo-Sotelo, O. Duque-Perez, R. Osornio-Rios, and M. Ibarra-Manzano, "Rotor unbalance and broken rotor bar detection in inverter-fed induction motors at start-up and steady-state regimes by high-resolution spectral analysis," *Electric Power Systems Research*, vol. 133, pp. 142–148, 2016.

[12] A. Naha, A. K. Samanta, A. Routray, and A. K. Deb, "A Method for Detecting Half-Broken Rotor Bar in Lightly Loaded Induction Motors Using Current," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 7, pp. 1614–1625, 2016.

[13] D. Matić, F. Kulić, M. Pineda-Sánchez, and I. Kamenko, "Support vector machine classifier for diagnosis in electrical machines: Application to broken bar," *Expert Systems with Applications*, vol. 39, no. 10, pp. 8681–8689, 2012.

[14] M. Abd-El-Malek, A. K. Abdelsalam, and O. E. Hassan, "Induction motor broken rotor bar fault location detection through envelope analysis of start-up current using Hilbert transform," *Mechanical Systems and Signal Processing*, vol. 93, pp. 332–350, 2017.

[15] J. Rangel-Magdaleno, H. Peregrina-Barreto, J. Ramirez-Cortes, and I. Cruz-Vega, "Hilbert spectrum analysis of induction motors for the detection of incipient broken rotor bars," *Measurement*, vol. 109, pp. 247–255, 2017.

[16] B. Bessam, A. Menacer, M. Boumehraz, and H. Cherif, "DWT and Hilbert Transform for Broken Rotor Bar Fault Diagnosis in Induction Machine at Low Load," *Energy Procedia*, vol. 74, pp. 1248–1257, 2015.

[17] R. A. Lizarraga-Morales, C. Rodriguez-Donate, E. Cabal-Yepez, M. Lopez-Ramirez, L. M. Ledesma-Carrillo, and E. R. Ferrucho-Alvarez, "Novel FPGA-based Methodology for Early Broken Rotor Bar Detection and Classification Through Homogeneity Estimation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1760–1769, 2017.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

Chapter 3

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

Mohammad Zawad Ali[1], *Student Member*, IEEE, Md Nasmus Sakib Khan Shabbir[1], *Student Member*, IEEE, Xiaodong Liang[1], *Senior Member*, IEEE, Yu Zhang[2], and Ting Hu[2]

[1]Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, Newfoundland, Canada.

[2]Department of Computer Science, Memorial University of Newfoundland, St. John's, Newfoundland, Canada.

A version of this chapter has been published in *IEEE Transactions on Industry Applications*, May/June 2019 regular issue. Mohammad Zawad Ali co-authored this paper under the supervision of Dr. Xiaodong Liang. Zawad's contributions are listed as follows:

- Performed literature searches required for background information of machine learning based fault diagnosis.

- Implemented hardware and performed experiments using two identical induction motors in the lab.

- Conducted signal processing and evaluated machine learning algorithms using experimented data.

- Examined the results and depicted the conclusion.

- Involved writing of the paper as the first author.

Dr. Xiaodong Liang provided the main ideas, set up the experimental plans, provided continuous technical guidance, checked the results, and modified the manuscript. Our research group member Md Nasmus Sakib Khan Shabbir participated in experiments, developed the curve fitting techniques and the

feature calculation formulas, wrote the relative part of the manuscript. Because machine learning content is the main technique in the paper, we have involved a graduate student, Yu Zhang, and a faculty member, Dr. Ting Hu, from the Department of Computer Science at Memorial University of Newfoundland. Yu Zhang reviewed and provided opinion for machine learning algorithms. Dr. Ting Hu reviewed the manuscript and provided expert opinion to improve the work.

In this chapter, the manuscript is presented with altered figure numbers, table numbers and reference formats in order to match the thesis formatting guidelines set out by Memorial University of Newfoundland.

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.

Keywords- Discrete wavelet transform, fault diagnosis, induction motors, machine learning, matching pursuit.

3.1 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 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 3.2, the machine learning based fault diagnosis approach using experimental data is proposed; Experimental set-up is provided in Section 3.3; in Section 3.4, 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 3.5; In Section 3.6, curve fitting equations are developed to calculate unknown features vs. motor loadings; conclusions are drawn in Section 3.7.

3.2 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. 3.1.

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.

5) Conduct classification for electrical and mechanical faults using chosen classifiers.

6) Develop curve fitting equations to calculate features vs. motor loadings.
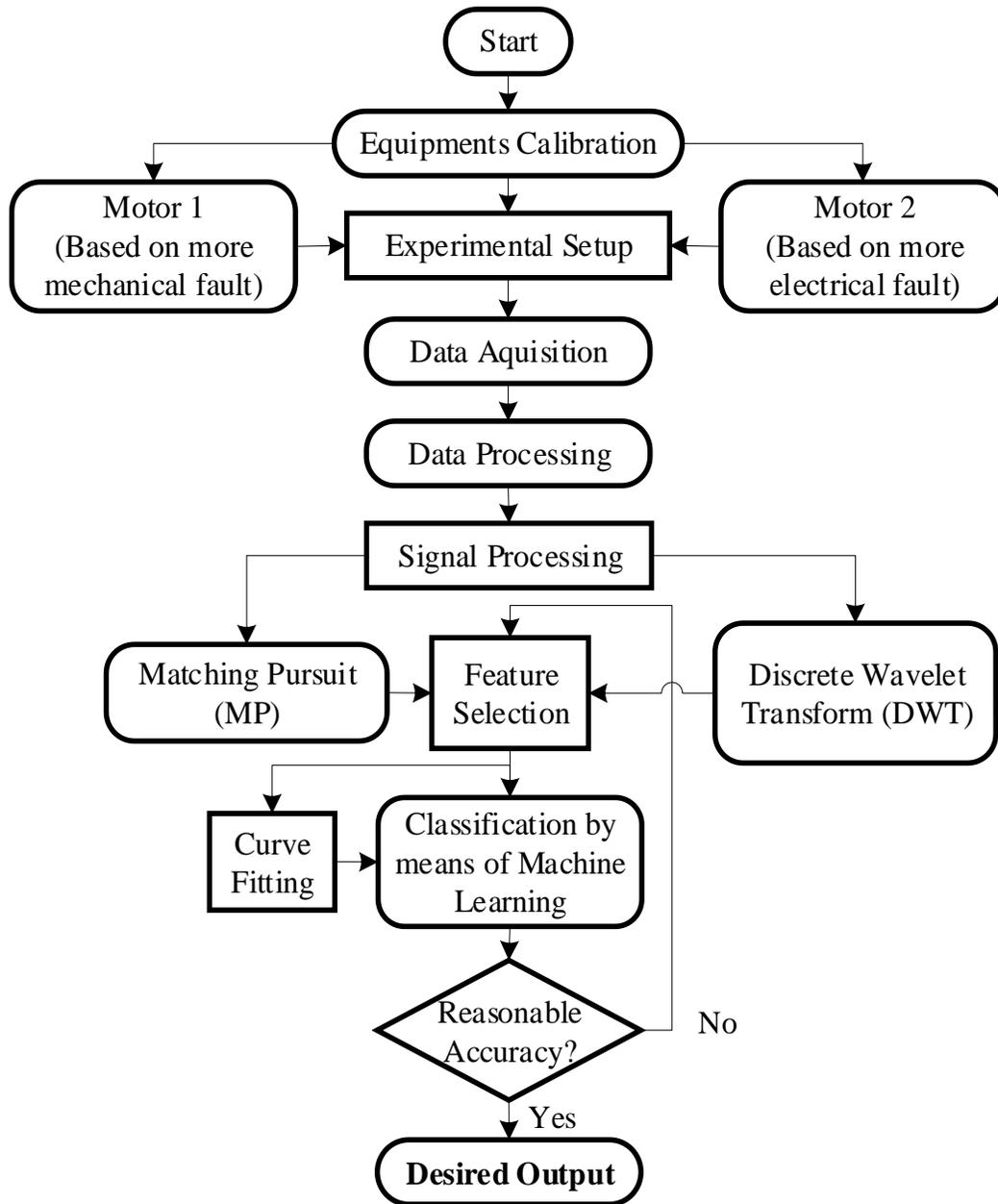
Fig. 3. 1. The flow chart of the proposed method.

3.3 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. 3.2.

(a)   (b)
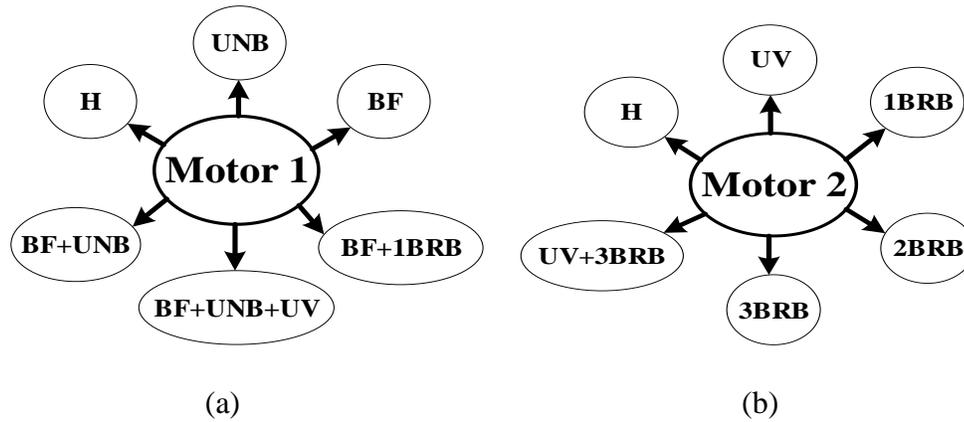
Fig. 3. 2. Experimental plan of the applied faults: (a) Motor 1; (b) Motor 2.

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.

In the experimental test bench (Fig. 3.3), 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.
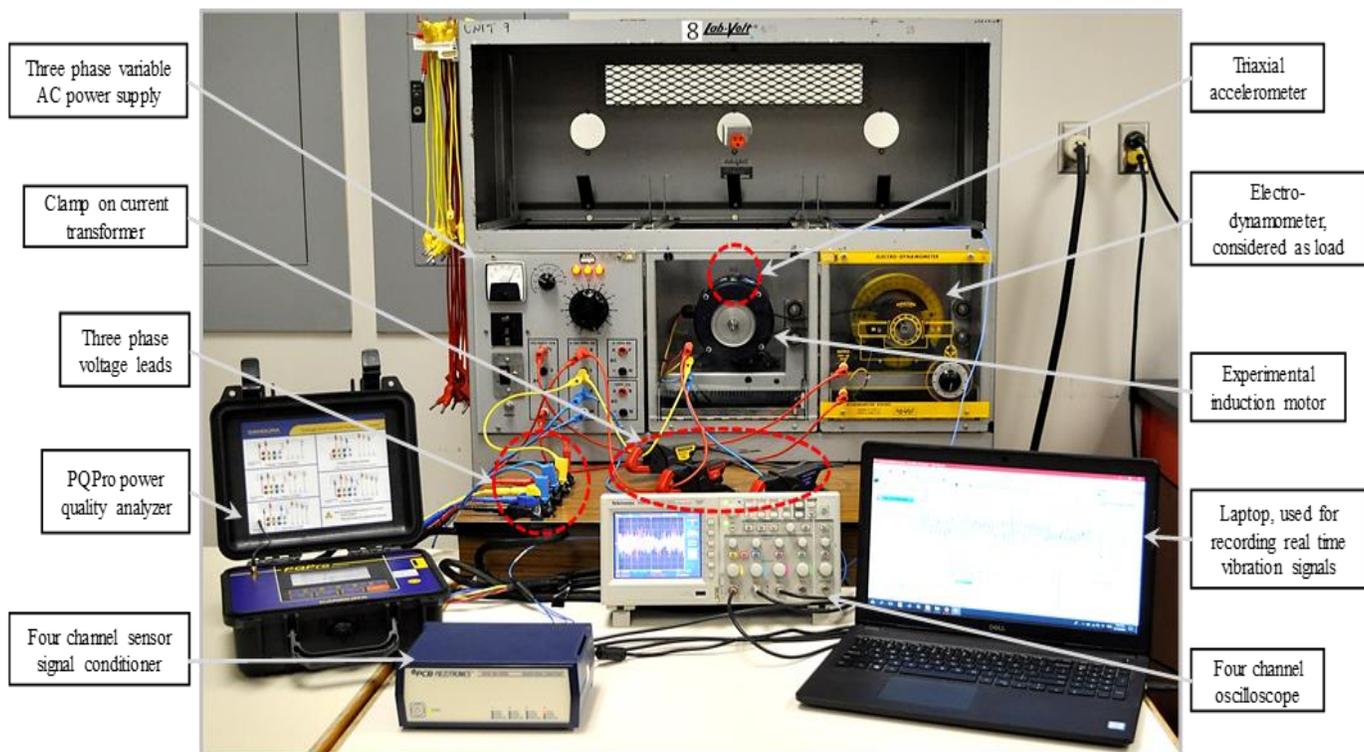
Fig. 3. 3. Experimental test bench used in this study.

As shown in Fig. 3.4, 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 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 (I1, I2, and I3) 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.
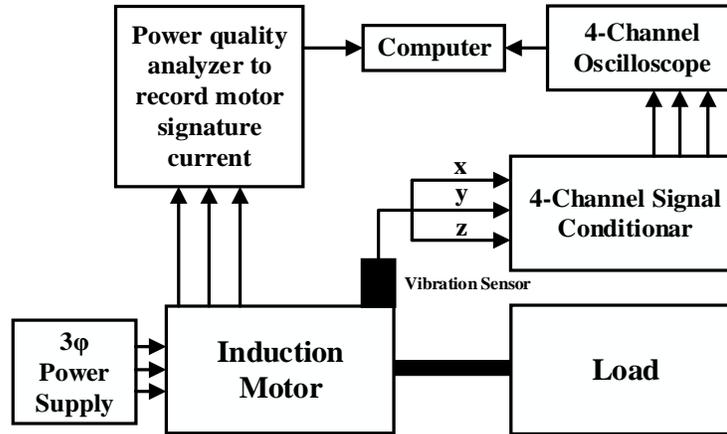
Fig. 3. 4. Experimental schematic diagram for the system set-up.

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. 3.5 (a)); two and three holes with 90° separation were drilled for two and three BRB faults, respectively (Figs. 3.5 (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. 3.5 (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. 3.5 (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)
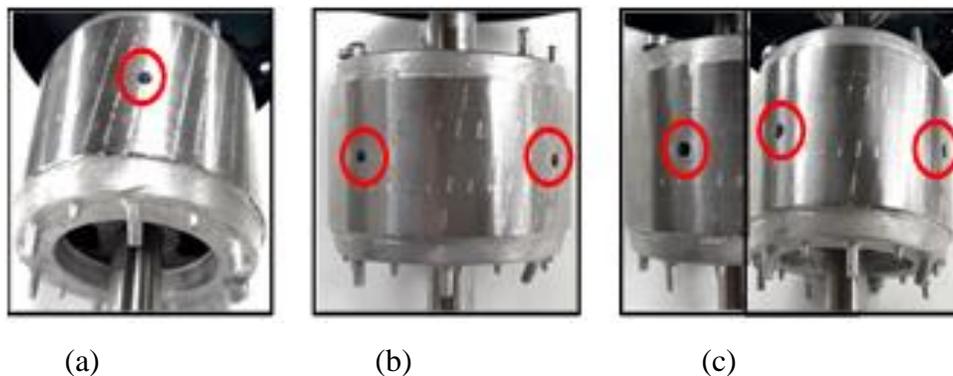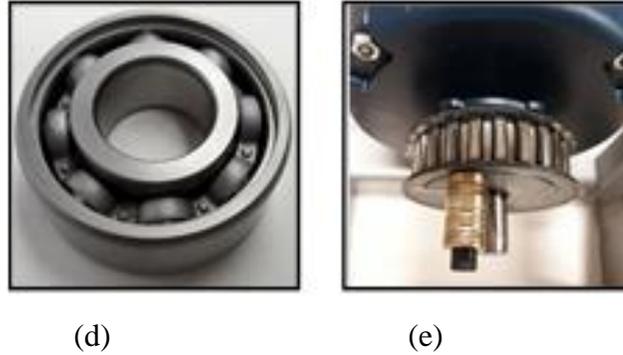
Fig. 3. 5. Implementation of different faults in the experimental test bench: (a) 1 BRB, (b) 2 BRB, (c) 3 BRB, (d) bearing fault – general roughness type, and (e) UNB condition.

3.4 Signal Processing for Feature Extraction

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

3.4.1. 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 (I2) 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. 3.6 using the current I2 and Fig. 3.7 using the z-axis vibration signal. In these figures, "3000" at the x-axis means 3000 sample points. In Figs. 3.6 (a) and 3.7 (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 sub dictionary; and 5) "cos", Cosine sub dictionary [32].

The dct and cos components are dominant in Fig. 3.6 (a), and the dct and sym4-lev5 components are dominant in Fig. 3.7 (a). By OMP processing, the approximated signals in Figs. 3.6 (c) and 3.7 (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.
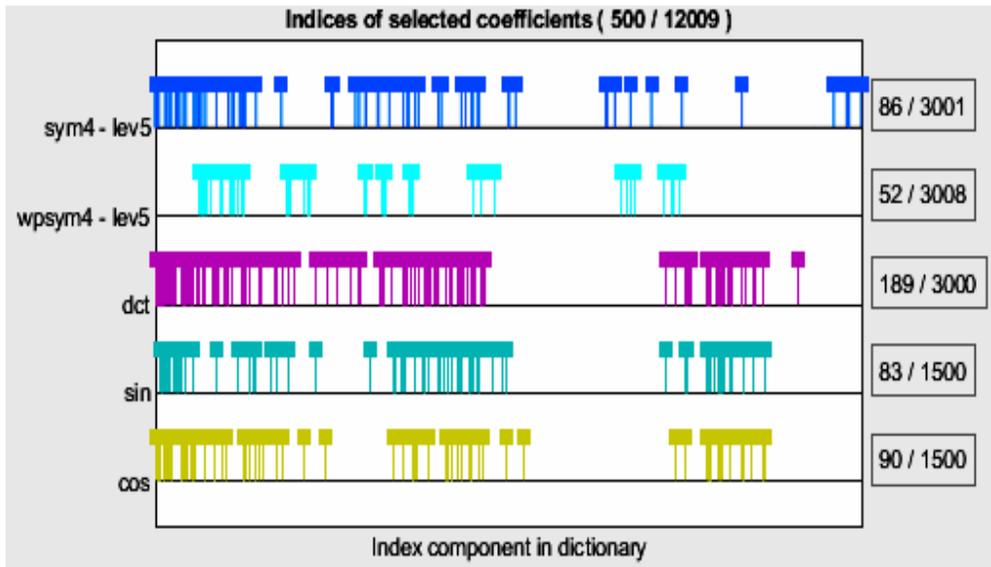
(a)
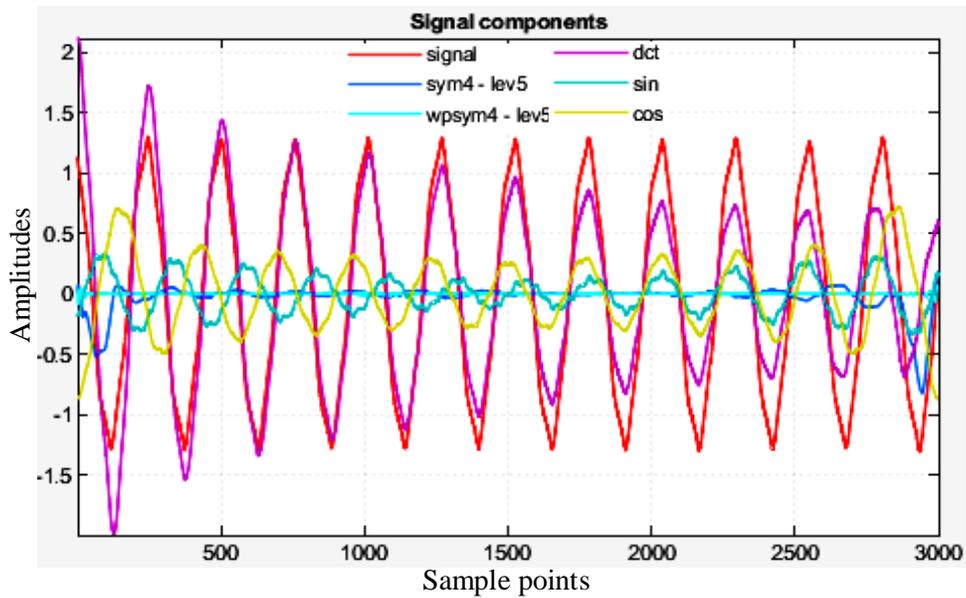
(b)
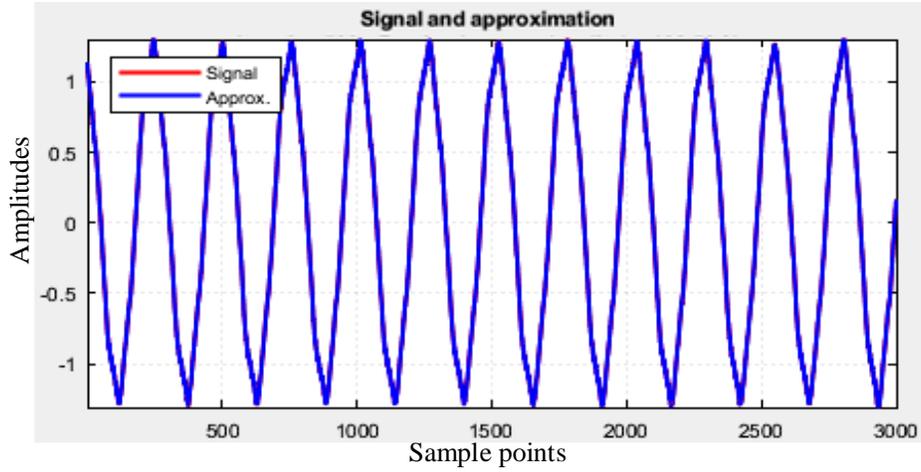
(c)

Fig. 3. 6. The stator current I2 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.
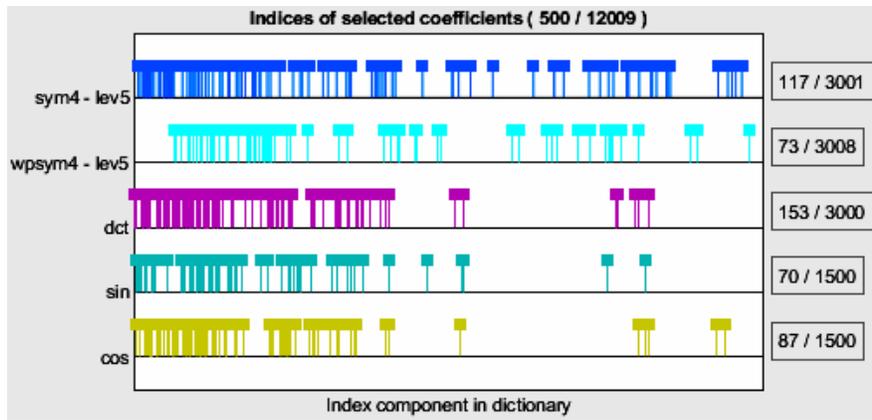
(a)
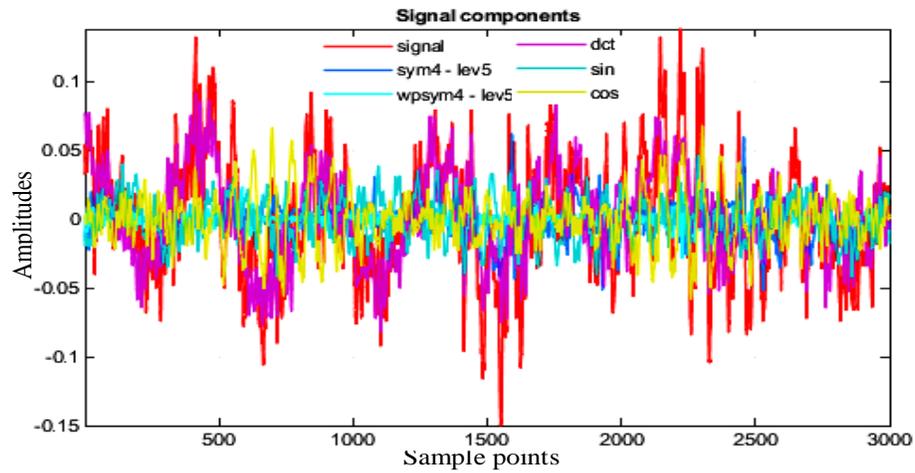
(b)
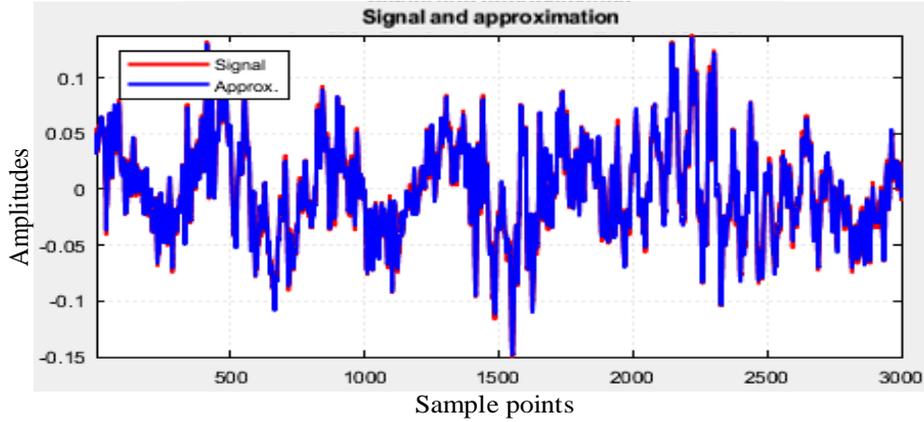
(c)

Fig. 3. 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 3.1 [33][34]. Table II shows a sample of features obtained using the current I2 for Motor 2 with a 1BRB fault at 100% loading. Every set of eight features, such as S1 in the first row of Table 3.2, is obtained by taking 3000 sample points from the current I2 and processed by the OMP. Other sets of features (from S2 to S7) are determined by taking sample points in a similar way. Fig. 3.8 shows one feature, Mean, for Motors 1 and 2 processed by the current I2 vs. motor loadings and different types of faults. Other features show similar patterns.

Table 3. 1: 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, $i = 1, 2, 3,\ldots$, N for N observations.
Median	$\text{med} = \frac{1}{2}(x_{(\lfloor(N+1)/2\rfloor)} + x_{(\lfloor N/2\rfloor+1)})$
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	$\text{Median\_AD} = median(|x_i - meidan(X)|)$
Mean Absolute Deviation	$\text{Mean\_AD} = \frac{1}{N}\sum_{i=1}^{N} |x_i - \mu_x|$

L1 norm	$\|L\|_1 = \sum_{i=1}^{N}|x_i|$, the sum of absolute values of its components, also known as one-norm, or mean norm
L2 norm	$\|L\|_2 = \sqrt{\sum_{i=1}^{N}|x_i|^2}$, 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)	$\|L\|_\infty = \max\{|x_i|: i = 1, 2, \ldots, n\}$, the maximum of absolute values of its components, also known as infinity norm, or uniform norm.

3.4.2. 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 I2 and z-axis vibration signals for Motor 2 with a 1 BRB fault at 100% motor loading are analyzed as shown in Figs. 3.9 and 3.10, 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 3.1 are also determined. Table 3.3 shows a sample of eight features processed using the stator current I2 for Motor 2 with a 1BRB fault at 100% loading.

Table 3. 2: A sample of Features using stator current I2 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

Table 3. 3: A sample of Features for Machine Learning using one phase stator current I2 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
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
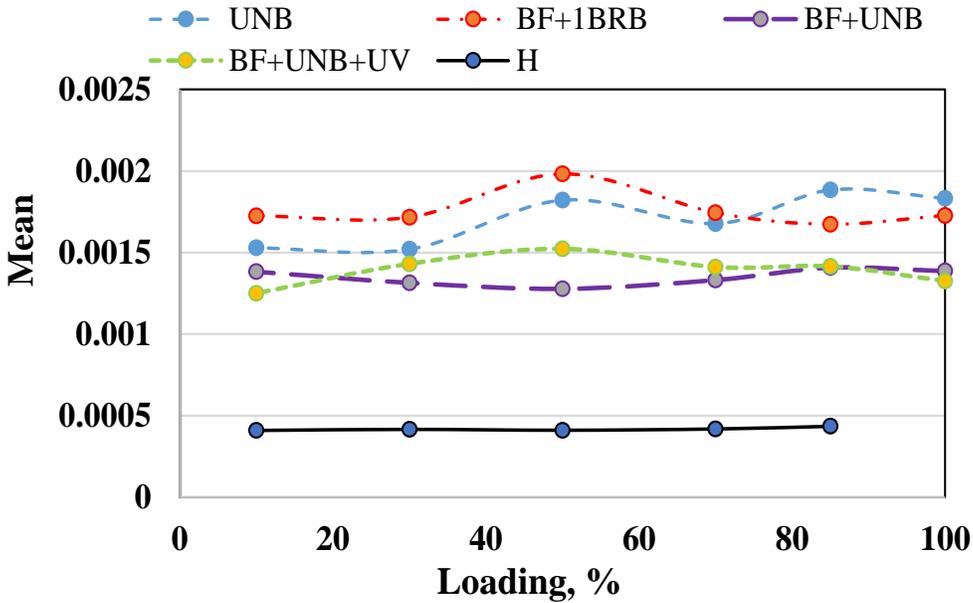
(a)
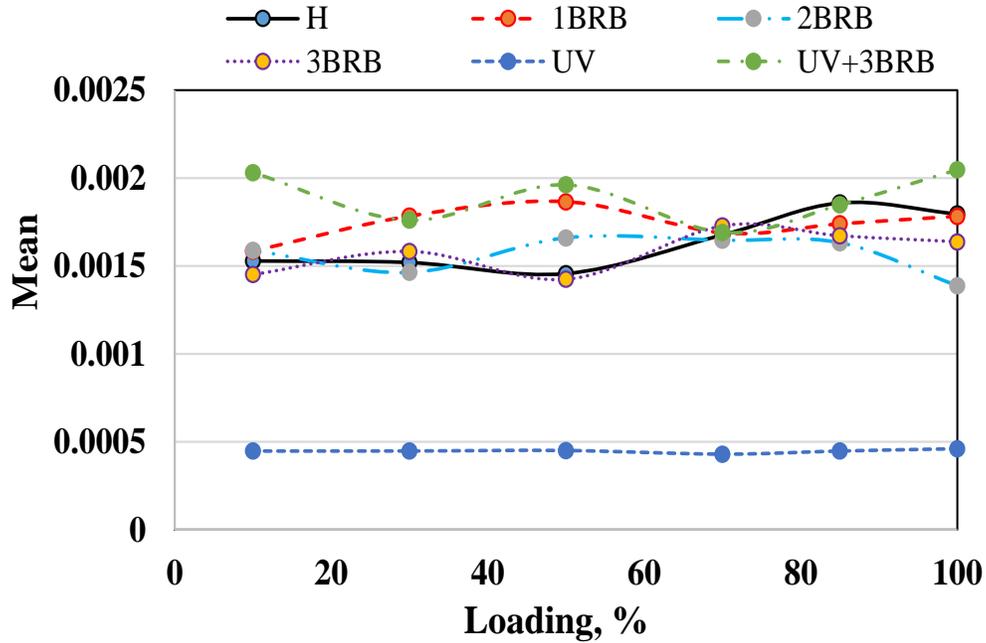
(b)

Fig. 3. 8. One feature, Mean, vs. motor loadings and different types of faults processed by OMP using the stator current I2: (a) Motor 1, and (b) Motor 2.
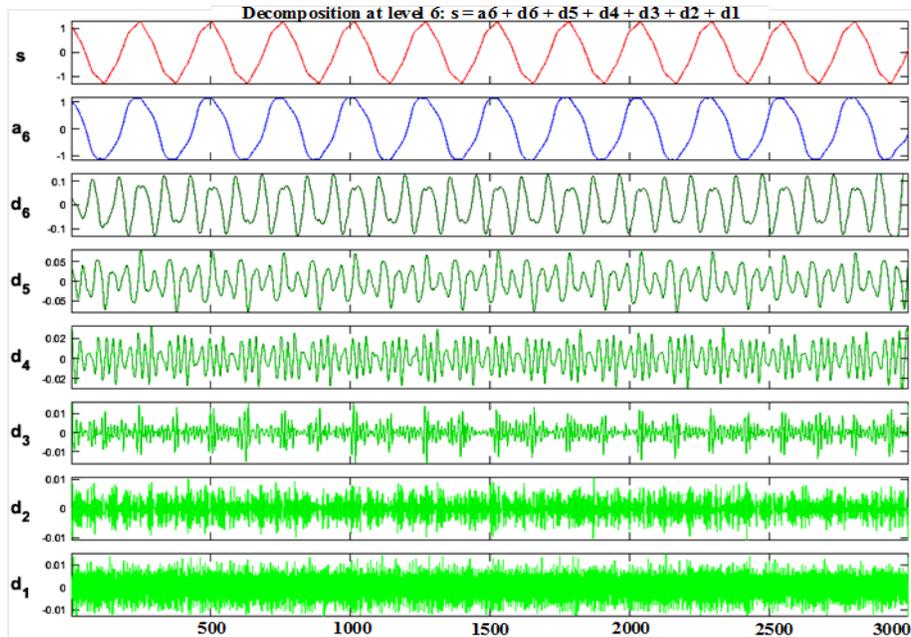
Fig. 3. 9. The processed one phase stator current signal I2 using DWT for Motor 2 under a 1 BRB fault and 100% loading condition.
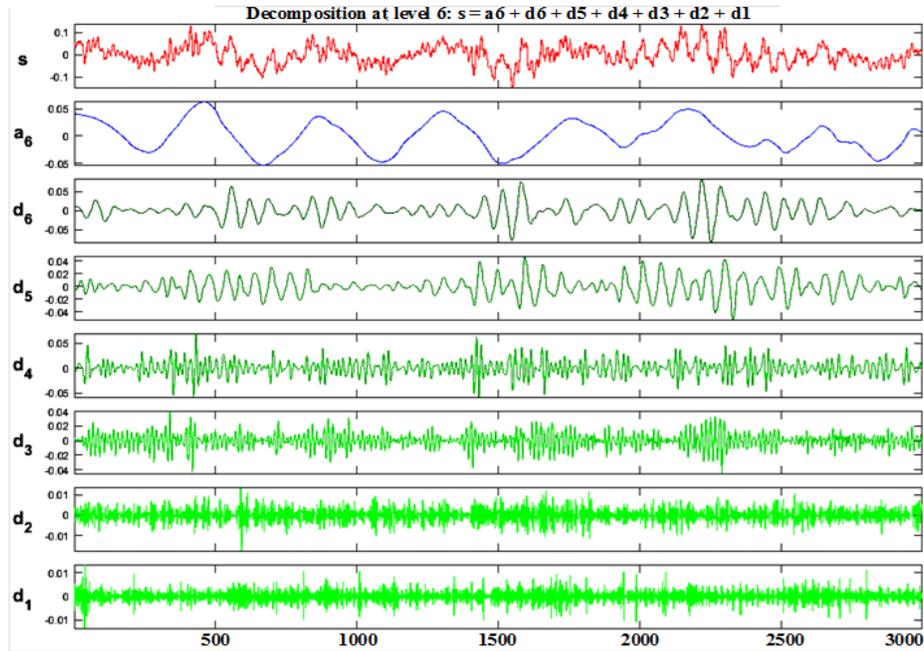
Fig. 3. 10. The processed z-axis vibration signal using DWT for Motor 2 under a 1 BRB fault and 100% loading condition.

3.5 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.

3.5.1 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 3.4 [40][41].

Table 3. 4: 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)	$k(x, y) = exp\left(-\dfrac{\|x - y\|^2}{2\sigma^2}\right)$ 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].

3.5.2 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 3.5 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 3. 5: 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 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.

3.5.3 Fault Diagnosis Results

The fault diagnosis accuracies for all faults of Motors 1 and 2 at 100% loading using the current I2 and z-axis vibration signal are shown in Figs. 3.11 and 3.12, respectively. In each graph, MP and DWT processing are compared. The data for Fig. 3.12 are also shown in Table 3.6.
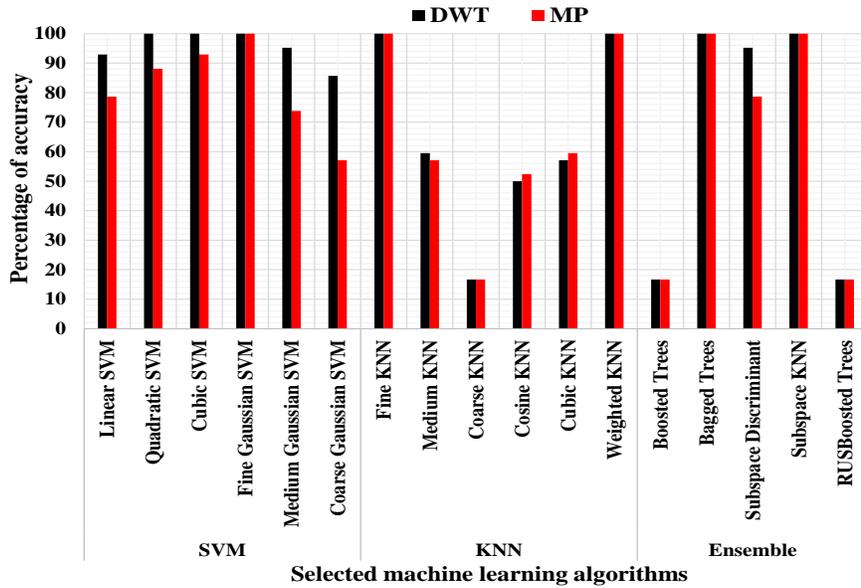
(a)
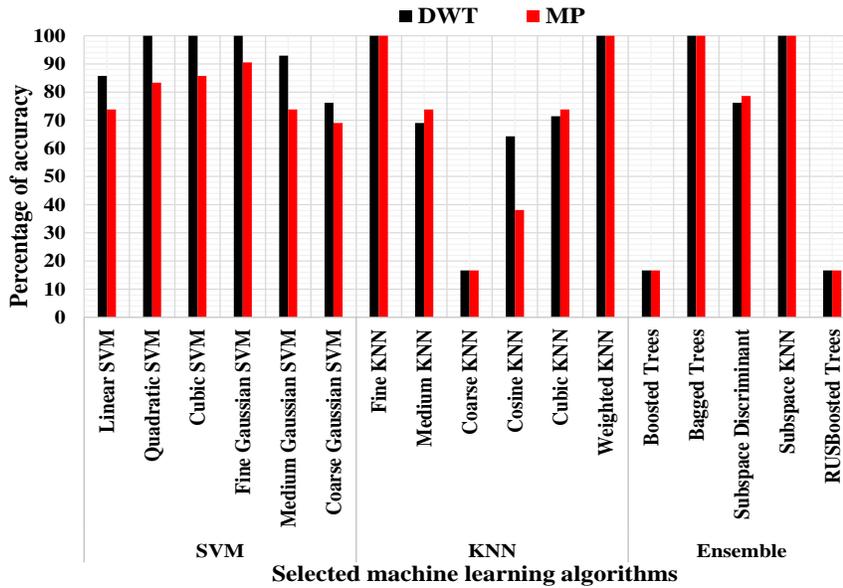
(b)

Fig. 3. 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.
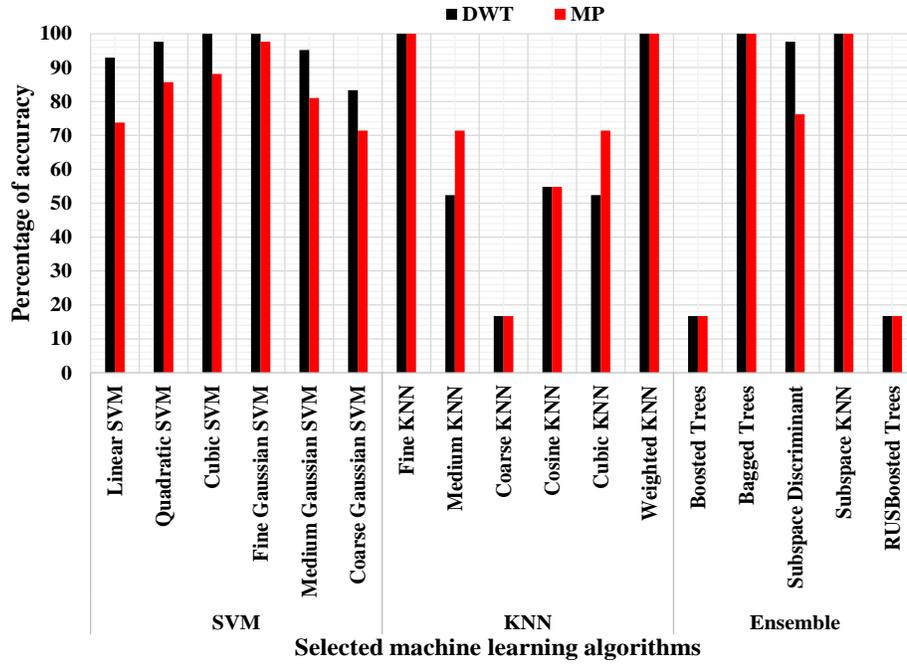
(a)
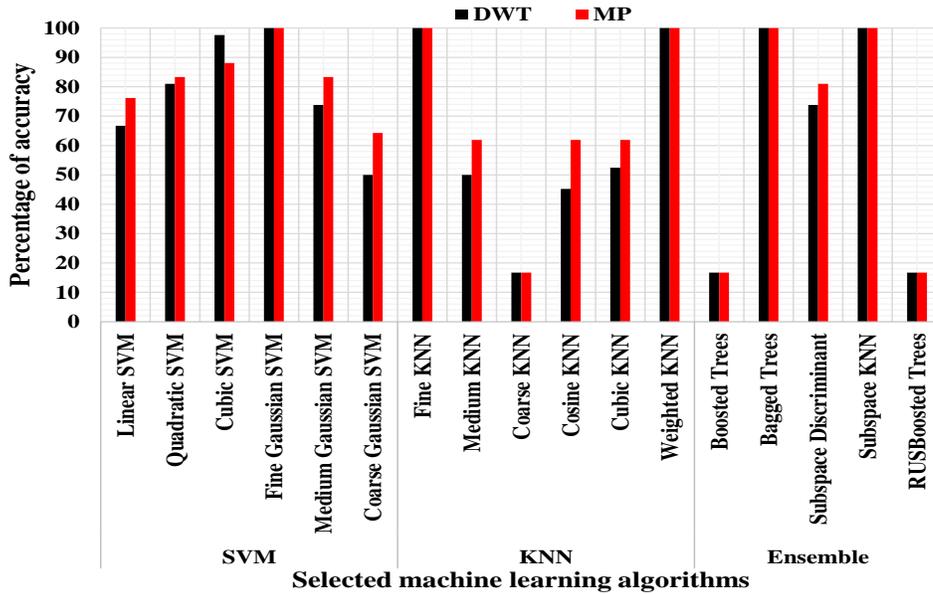
(b)

Fig. 3. 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 3. 6: 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
	Bagged Trees	**100**	**100**	**100**	**100**
	Subspace Discriminant	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:

$$\text{Accuracy} = \frac{TP}{TP+FN} \tag{1}$$

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]:

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

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

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. 3.13 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 I2 signal.
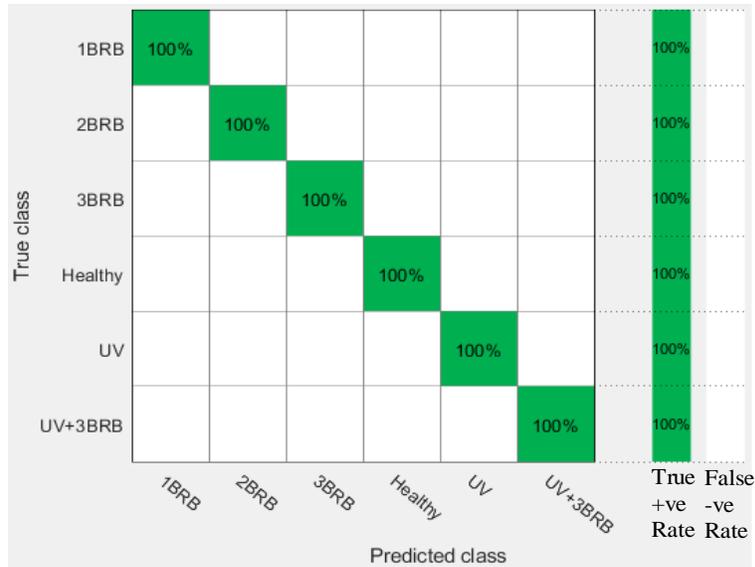
(a)
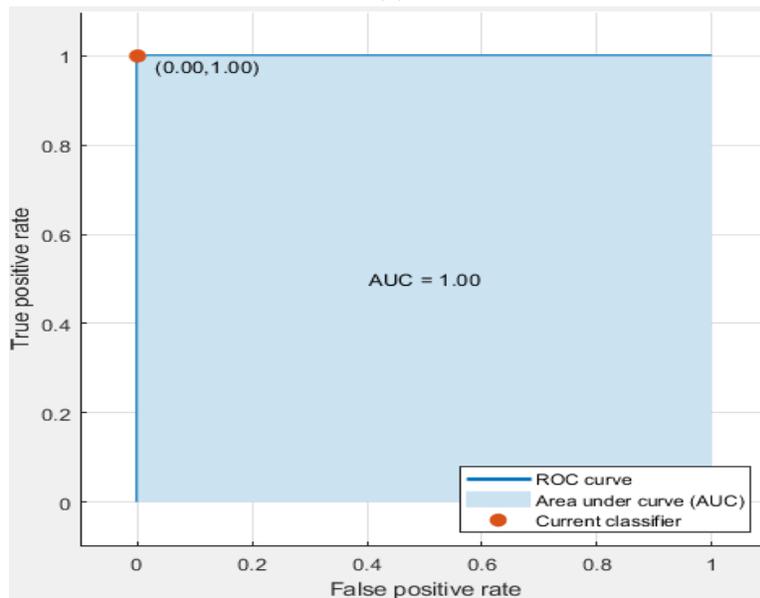
(b)

Fig. 3. 13. 100% classification accuracy obtained by Fine KNN for Motor 2 at 100% loading using the current I2: (a) confusion matrix; (b) ROC curve.

3.5.4 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. 3.11 and 3.12 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.

3.5.5 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 3.7. 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 3. 7: 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

3.6 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.

3.6.1 Curve Fitting Method

The purpose of the proposed curve fitting technique is to find statistical features for untested loading conditions under motor healthy and faulty cases. After getting all features for untested loadings through curve fitting, the same method used for the tested loading conditions is followed for fault classification.

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 3.8 shows regression models along with their R-square values for Motor 2 with a 1BRB fault processed by MP using the stator current I2. 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 3.9 with all errors less than 8%, which further validates the accuracy of the fitting equations. Fig. 3.14 shows the graphs of the eight features vs. the motor loading using the stator current I2 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 3. 8: Regression models for features using stator current I2 processed by MP for Motor 2, 1 BRB fault

Feature Name	Equation	R-square Values
Mean	$y = -2E\text{-}07x^2 + 2E\text{-}05x + 0.0013$	0.9512
Median	$y = -1E\text{-}07x^2 + 2E\text{-}05x + 0.0011$	0.9197
Standard Deviation	$y = -1E\text{-}07x^2 + 1E\text{-}05x + 0.001$	0.9897
Median Absolute Value	$y = -8E\text{-}08x^2 + 9E\text{-}06x + 0.0006$	0.9168
Mean Absolute Value	$y = -8E\text{-}08x^2 + 1E\text{-}05x + 0.0008$	0.9700

L1 Norm	$y = -0.0005x^2 + 0.0549x + 3.86$	0.9512
L2 Norm	$y = -1E\text{-}05x^2 + 0.0012x + 0.0898$	0.9695
Maximum Norm	$y = -6E\text{-}07x^2 + 7E\text{-}05x + 0.006$	0.6482

Table 3. 9: Relative errors between experimental based data and calculated data (for Motor 2, 1 BRB fault, stator current I2)

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

Similarly, curve fitting can be applied to vibration signal to obtain features of a new motor loading for a fault. Table 3.10 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 3.11. Fig. 3.15 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 3. 10: 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 3. 11: 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
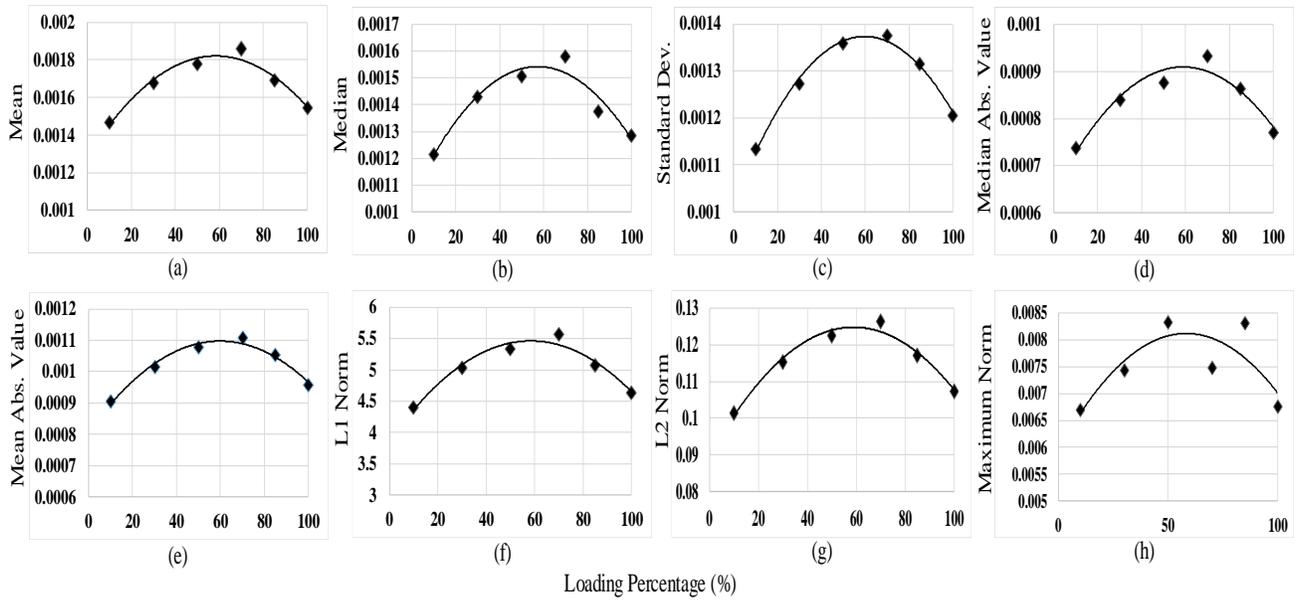
Fig. 3. 14. Curve fitting results for features of Motor 2 with a 1BRB fault using the stator current I2: (a) mean, (b) median, (c) standard deviation, (d) median absolute value, (e) mean absolute value, (f) L1 norm, (g) L2 norm, and (h) maximum norm.
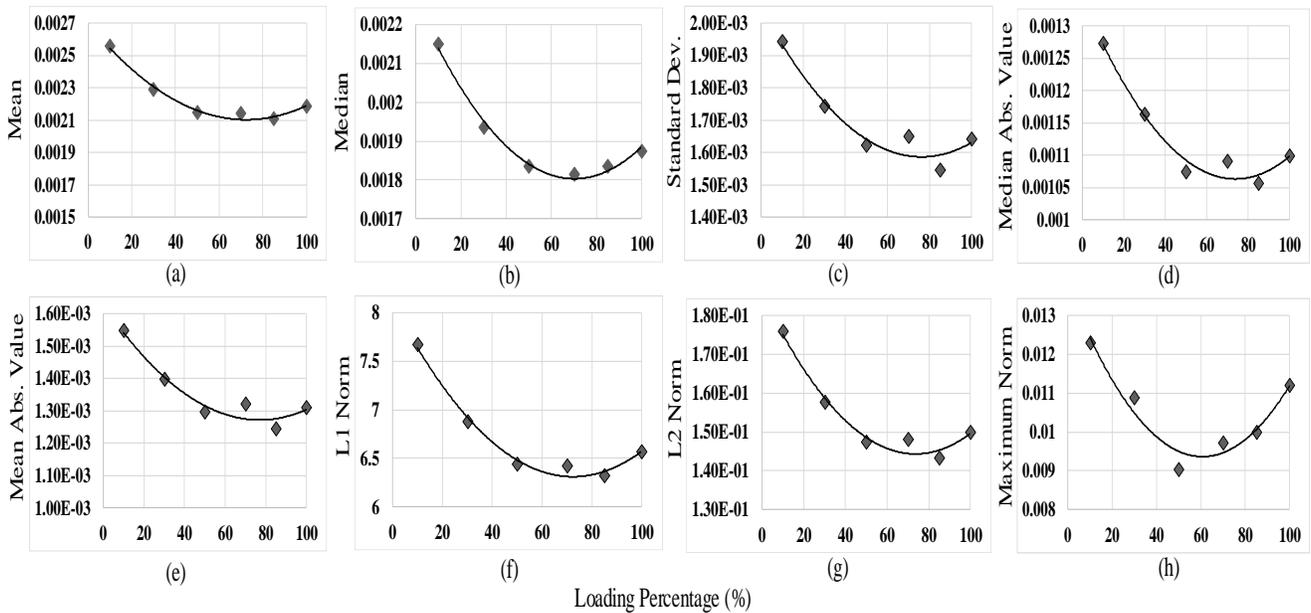
Fig. 3. 15. Curve fitting results for features of Motor 2, 1BRB fault using the z-axis vibration signal: (a) mean, (b) median, (c) standard deviation, (d) median absolute value, (e) mean absolute value, (f) L1 norm, (g) L2 norm, and (h) maximum norm.

3.6.2 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. 3.16 shows fault classification accuracy for the three loadings for Motor 2 with the current I2. Curve fitting equations offer effective calculation of unknown features for various motor loadings.
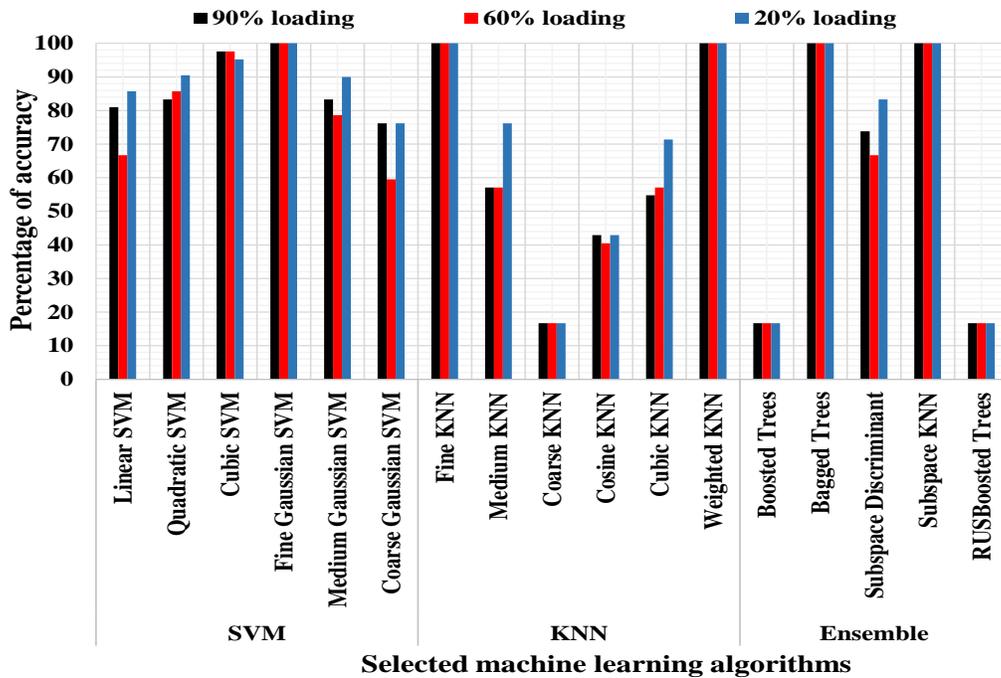
Fig. 3. 16. Classification accuracy for all faults using features calculated by curve fitting equations for three loadings (90%, 60% and 20%) that has never been tested by experiments (Motor 2, the stator current I_2).

3.7 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.

Chapter 4

Induction Motor Fault Diagnosis Using Discrete Wavelet Transform

Mohammad Zawad Ali, *Student Member*, IEEE, and Xiaodong Liang, *Senior Member*, IEEE

Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, Newfoundland, Canada.

A version of this chapter has been published on the Proceedings of 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE). Mohammad Zawad Ali developed this work under the supervision of Dr. Xiaodong Liang. Zawad's contributions in this paper are listed as follows:

- Performed literature searches required for background information on induction motors fault diagnosis.

- Using part of the experimental data obtained in Chapter 3 in this paper to develop a threshold value based fault diagnosis.

- Examined the results and discussed the findings.

- Wrote the paper.

Dr. Xiaodong Liang provided continuous technical guidance, checked the results, reviewed the manuscript, provided important suggestions to accomplish the work, and modified the final version of the manuscript.

In this chapter, the manuscript is presented with altered figure numbers, table numbers and reference formats in order to match the thesis formatting guidelines set out by Memorial University of Newfoundland.

Abstract- In this paper, a general methodology is constructed by using experimentally measured stator current signals under full load condition of an induction motor. The measured stator current data for various single- and multi-electrical faults of the induction motor are investigated to obtain signatures for fault

diagnosis. In this study, the discrete wavelet transform (DWT) is chosen for signal processing. The threshold and energy values at each decomposition level for the DWT analysis are evaluated. The threshold values appear to be more consistent than the energy values at different data windows of the measurement data, and thus, the threshold can serve as a reliable parameter for fault diagnosis.

Keywords- Fault diagnosis, discrete wavelet transform, induction motors, broken rotor bar.

4.1 Introduction

Induction motors are workhorse for our modern industry. Condition monitoring and fault diagnosis of induction motors play an important role to maintain reliable and smooth operation of industrial processes [1]-[4]. An unexpected motor breakdown may interrupt the workflow and reduce the motor drive efficiency. Therefore, condition monitoring plays a significant role to maintain reliable and smooth operation in industrial processes [1].

The motor current signature analysis (MCSA) is the most commonly used fault diagnosis method [5]-[7]. Advanced signal processing techniques are reported in the literature as a critical step for fault diagnosis. These techniques include wavelet transform [1][6]-[9], Multiple Signal Classification (MUSIC) method [10]-[12], Hilbert Transform [13]-[15], or hybrid techniques, such as combining Wavelet and Hilbert transforms with a linear discrimination method [16], and homogeneity analysis with Gaussian probability density function [17].

In this paper, the Discrete Wavelet Transform (DWT) [6][7][18] is implemented for signal processing using the stator current signals measured in a lab using a 0.25 HP squirrel-cage induction motor for the fault diagnosis purpose. The suitability of the DWT method is assessed by the threshold value of each decomposition level and the energy of each detail level. A robust fault diagnosis method is proposed for classifying various faults of induction motors based on the DWT processing results.

The paper is detailed as follows: in Section 4.2, the proposed approach is demonstrated through a recommended 3-step procedure, and the experimental set-up for testing the 0.25 HP induction motor is explained; a brief explanation of the DWT method is provided in Section 4.3; in Section 4.4, the signal processing results using the DWT are demonstrated and analyzed, the fault diagnosis criteria are summarized; conclusions are drawn in Section 4.5.

4.2 The proposed Method and Experimental Test Bench

In this paper, a fault diagnosis method for induction motors using the stator current signal is proposed. The main idea is shown in Fig. 4.1. There are three critical steps involved in the method: 1) The experiment is conducted for healthy and several faulty conditions, faults are prepared by damaging the motor physically; 2) After equipment calibration and experimental setup, the motor stator current is measured for each condition; 3) Signal processing using the DWT analysis to extract fault signatures. In DWT, decomposition levels are specified, the threshold and energy values associated with each decomposition level are calculated, which will be used for fault diagnosis.
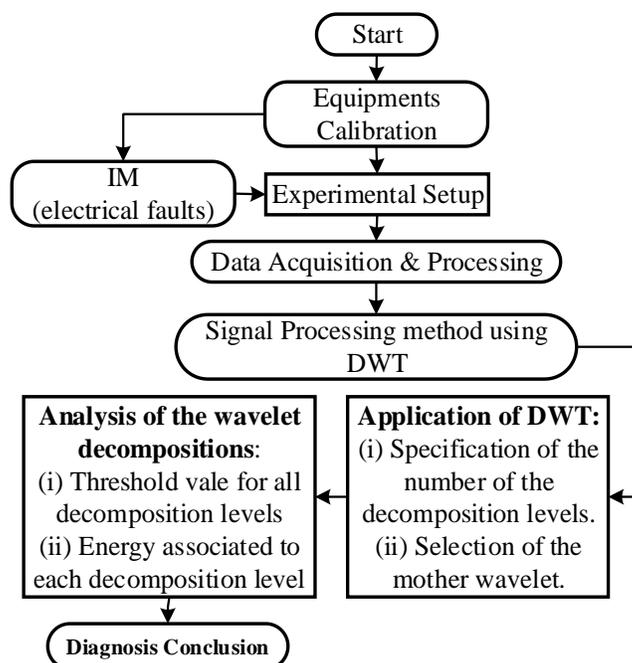
Fig. 4. 1. The flow chart of the proposed method.

In this study, a 4-pole, ¼ HP, 208-230/460V squirrel cage induction motor is used in experiments. The healthy and faulty conditions of the motor are shown in Fig. 4.2. The experiments were conducted on: (a) healthy motor (H); (b) unbalance voltage (UV) condition; (c) one broken rotor bar (BRB) fault; (d) two BRB fault; (e) three BRB fault; and (f) a multi-fault by combining UV and 3 BRB faults. The UV condition is formed by adding an extra resistance on one phase of the power supply. The BRB faults are produced by drilling a hole (4.2 mm diameter and 18 mm depth) on the rotor bar. One hole is drilled for a 1 BRB fault; two and three holes with 90° separation are drilled for 2 BRB and 3BRB faults as shown in Fig. 4.3. The induction motor is connected directly to a three phase power supply, and the motor shaft is coupled through a belt pulley with a dynamometer as the load. At the rated speed, the full load of the motor is measured corresponding to a 7 pound force inch (lbf-in) torque. The three-phase stator currents are recorded using an 8-channel power quality analyzer (PQPro by CANDURA instrument).
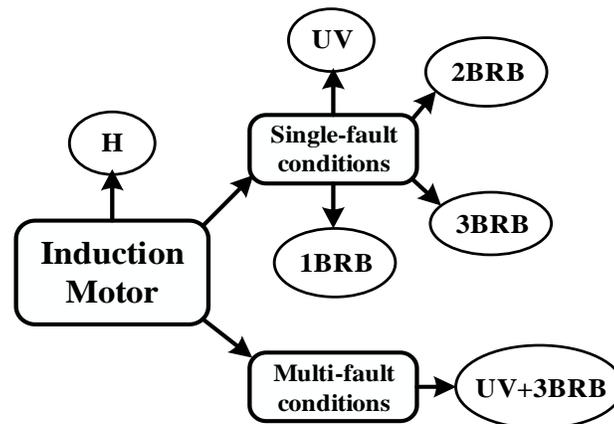
Fig. 4. 2. Detailed experiment plan for healthy and faulty conditions.

(a)   (b)   (c)
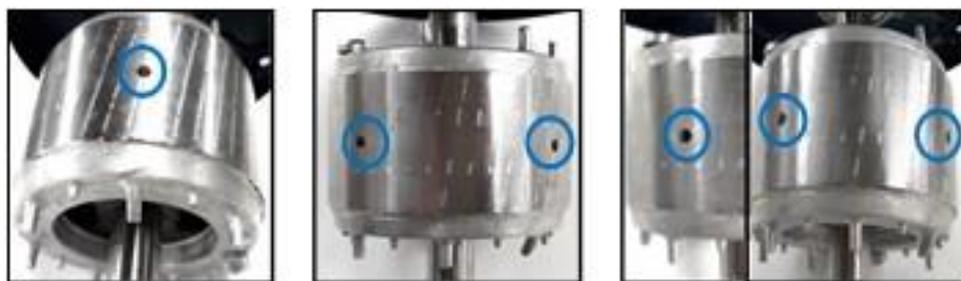
Fig. 4. 3. Motors with BRB faults: (a) 1 BRB, (b) 2 BRB, and (c) 3 BRB.

4.3 Signal Processing Approaches

The DWT offers an effective analysis for time-frequency representation of a non-stationary signal [6][7]. Orthogonal wavelets such as Daubechies wavelet series are used to decompose the signal into several frequency bands [8]. Through the DWT, an original signal is decomposed into several batches of wavelet signals, each contains the original signal's information within a certain frequency band. The number of decomposition is known as levels. The decomposition can be implemented using filtering and down sampling process as shown in Fig. 4.4. the operation is repeated until the signal is decomposed to the preferred level.
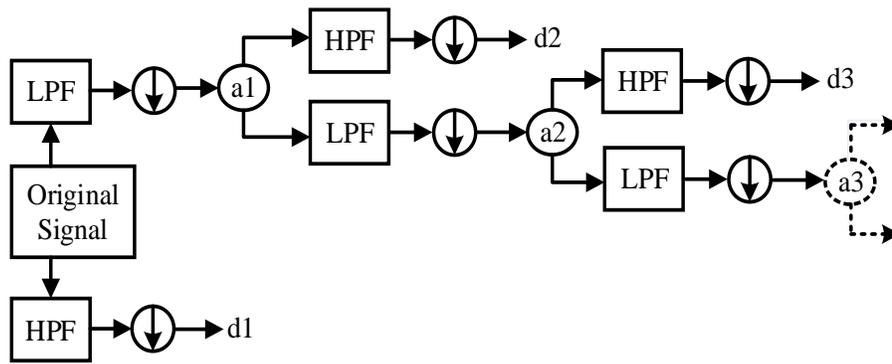
Fig. 4. 4. Sequence of signal decomposition process into approximations and details by DWT.

At each decomposition stage, two coefficients can be determined: 1) approximation coefficient aj, which are acquired from the low pass filter (LPF), containing low frequency signal components; 2) detail coefficient dj, which are acquired from the high pass filter (HPF), containing high frequency signal components [6][8], where j is the decomposition level. The selection criteria for the appropriate mother wavelet and the number of the decomposition levels are key steps in the DWT analysis. In this study, Daubechies-44 (db44) wavelet is selected as the mother wavelet because it provides a more precise detail signal with lower harmonics. The total number of decomposition level N_{Ls} can be calculated as follows [7][9][20]:

$$N_{LS} = int\left(\frac{log\left(\frac{f_S}{f}\right)}{log\left(2\right)}\right) + 2 \qquad (1)$$

Where, f_s is the sampling frequency for the captured signal (In this study, f_s is approximately equal to 15.5 kHz); f is the fundamental frequency (60 Hz); 2 means that two more additional decomposition levels are suitable. Eq. (1) leads to 10 level decompositions. The detail coefficient d_j and the approximations coefficient a_j have the following frequency bands [9]:

$$f_{dj} \in \left[\left(\frac{f_S}{2^{(j+1)}}\right) \rightarrow \left(\frac{f_S}{2^j}\right)\right] Hz \qquad (2)$$

$$f_{aj} \in \left[0 \rightarrow \left(\frac{f_S}{2^{(j+1)}}\right)\right] Hz \qquad (3)$$

In this paper, the measured stator current signal is processed using DWT, frequency bands for each approximation and detail signals from levels 1 to 10 is determined and tabulated in Table 4.1 using (1)-(3).

Table 4. 1: Frequency Bands for Multi-levels Decomposition Obtained by DWT

Levels		Approximation signals, a_j (Hz)		Detail signals, d_j (Hz)
j = 1	a_1	0-3850	d_1	3850-7700
j = 2	a_2	0-1925	d_2	1925-3850
j = 3	a_3	0-962.5	d_3	962.5-1925
j = 4	a_4	0-481.25	d_4	481.25-962.5
j = 5	a_5	0-240.625	d_5	240.625-481.25
j = 6	a_6	0-120.3125	d_6	120.3125-240.625
j = 7	a_7	0-60.1563	d_7	60.1563-120.3125
j = 8	a_8	0-30.0781	d_8	30.0781-60.1563
j = 9	a_9	0-15.0391	d_9	15.0391-30.0781
j = 10	a_{10}	0-7.5195	d_{10}	7.5195-15.0391

4.4 Signal Processing Results Using DWT

In this paper, the MATLAB wavelet toolbox is used for the DWT analysis. db44 is the mother wavelet under the 10th level decompositions. The analysis is conducted using four data windows from the measured stator current signal, each data window consists of 4000 sample points. Data windows 1-4 are [68.2223 s, 68.4824 s], [80 s, 80.26 s], [72 s, 72.261 s], and [77.25 s, 77.51 s], respectively.

As an example, the motor stator current with 1 BRB fault under 100% loading is analyzed using the data window 1. Fig. 4.5 shows the original signal along with the details plotted for levels 1-10. It is observed that the activity in the detail signals reduces drastically as the scale or decomposition level increases. Based on the level 1 detail and ignoring the rest of the levels, the aim here is to retain sharp changes and get rid of the noise, which can be done by scaling detail coefficients by a threshold. The universal threshold (UT) technique is followed to compute the threshold as follows [21][22]:

$$UT = \frac{\sqrt{2*log\big(length(x)\big)*Median\ (abs(D))}}{0.6745} \tag{4}$$

Where, x is the signal, and D is the set of first level detail coefficients. Later, the threshold value is determined for all other detail levels. The computation of the threshold value can be done by soft or hard thresholding operations, in both cases, coefficients with the magnitude less than the threshold are set to zero. In this paper, the soft thresholding operation is considered, and the coefficients in magnitude greater than the threshold are shrunk towards zero. Based on all motor conditions, the threshold values for all decomposition levels are determined as shown in Fig. 4.6 using four data windows, which can be considered as a fault indicator. The energy associated with each decomposition level is evaluated to see if it can be used in fault diagnosis. The energy of each frequency band can be calculated by [7][20]

$$E_j = \sum_{n=1}^{N}\big|d_j(n)\big|^2 \tag{5}$$

Where, N is the number of samples, d_j is the detail signal at the level j. Based on (5), the energy for each frequency band is calculated as shown in Fig. 4.7 using four data windows. It is found that the changes are obvious between levels d6 to d9. At level d8, different motor fault conditions can be determined efficiently.

In this study, the main tasks for the DWT is to calculate the threshold value for each decomposition level and the energy for each frequency bandwidth. It is observed using the four data windows that the changes

of both threshold and energy occur between d7 and d9 decomposition levels. The summarized threshold and energy values at d8 are provided in Table 4.2. The tendency of the changes is consistent for threshold and energy for healthy and faulty conditions of the motor using both data windows. The changes are more pronounced between d7 and d9, especially the energy reaches the highest values at d8. However, the threshold appears to be more stable for the four data windows, with very similar shape and magnitude values, while the shape and magnitude of the energy varies quite significantly. Therefore, the threshold is considered as a more reliable fault detection criteria.
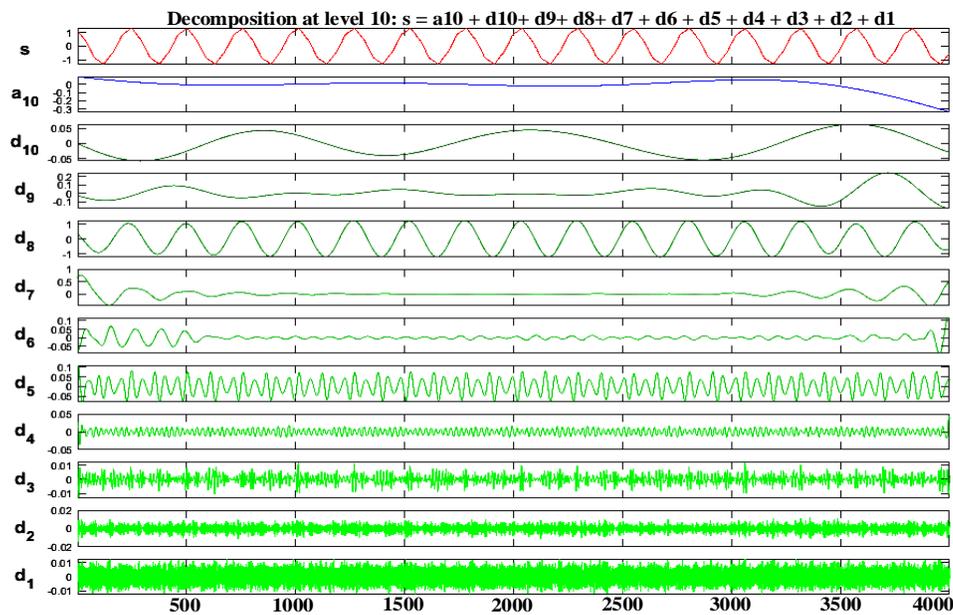
Fig. 4. 5. The processed stator current signal I2 of the motor at 100% loading, 1 BRB condition using DWT.
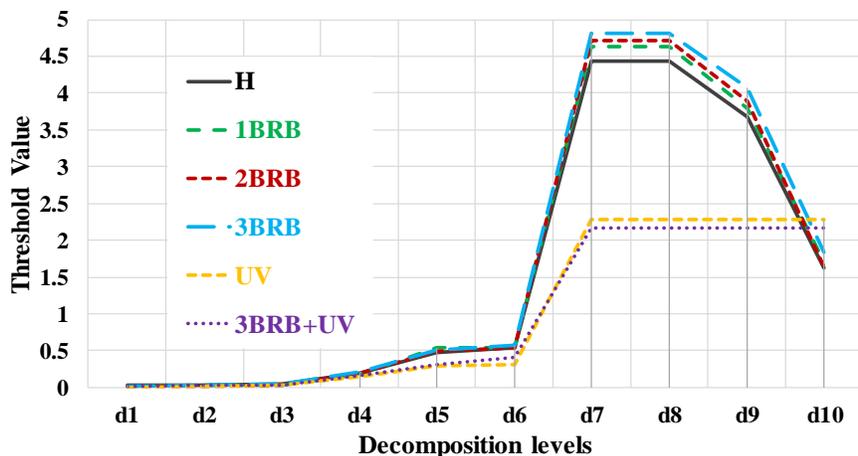
(a)
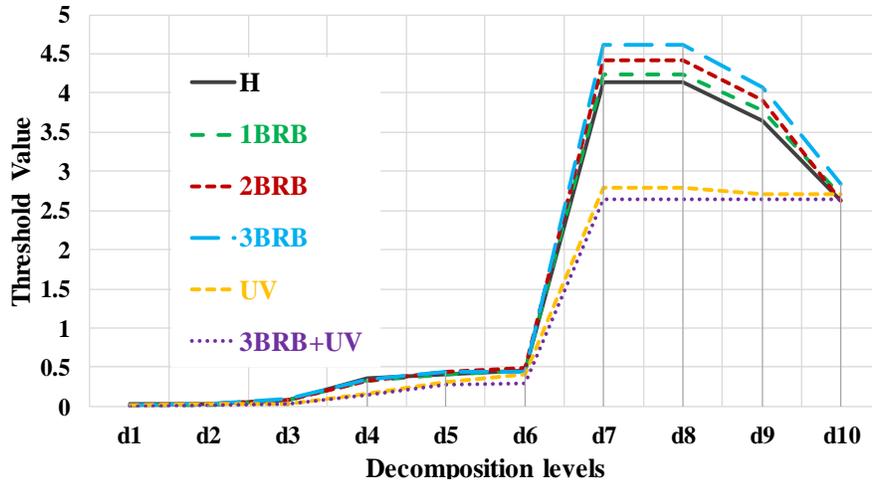
(b)
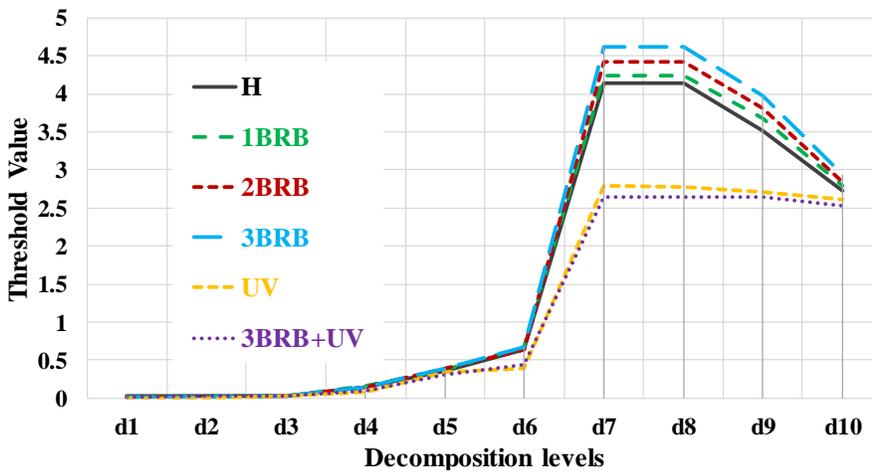
(c)
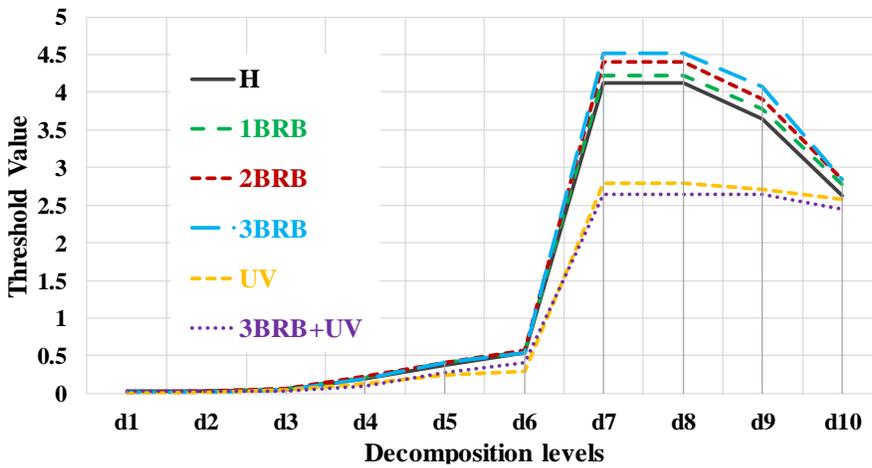
(d)

Fig. 4. 6. Threshold values for all decomposition levels using four different data windows in measured stator current: (a) data window 1; (b) data window 2; (c) data window 3; and (d) data window 4.

(a)

(b)

(c)

(d)
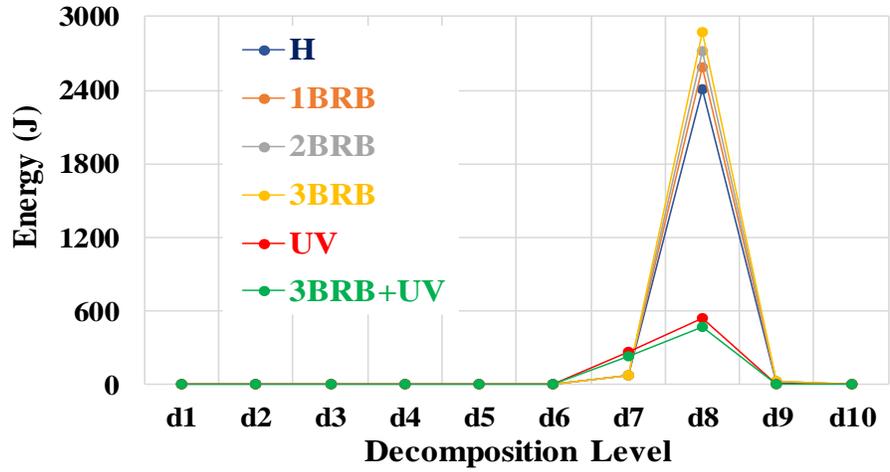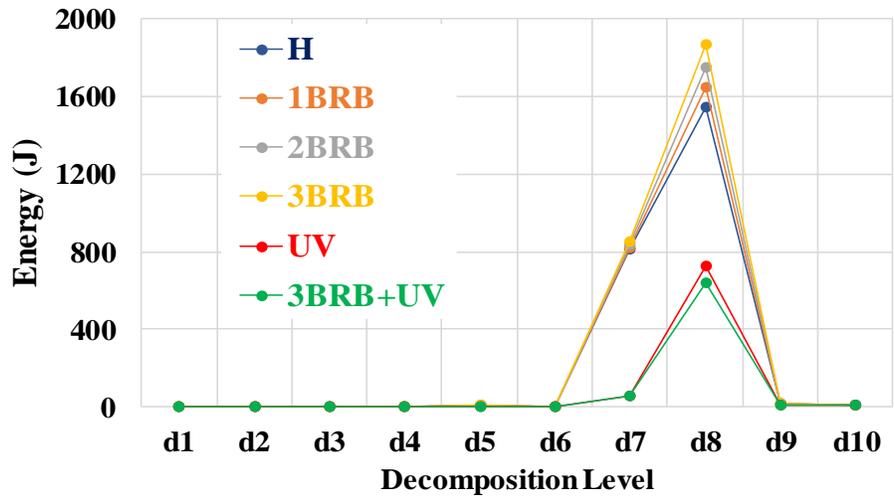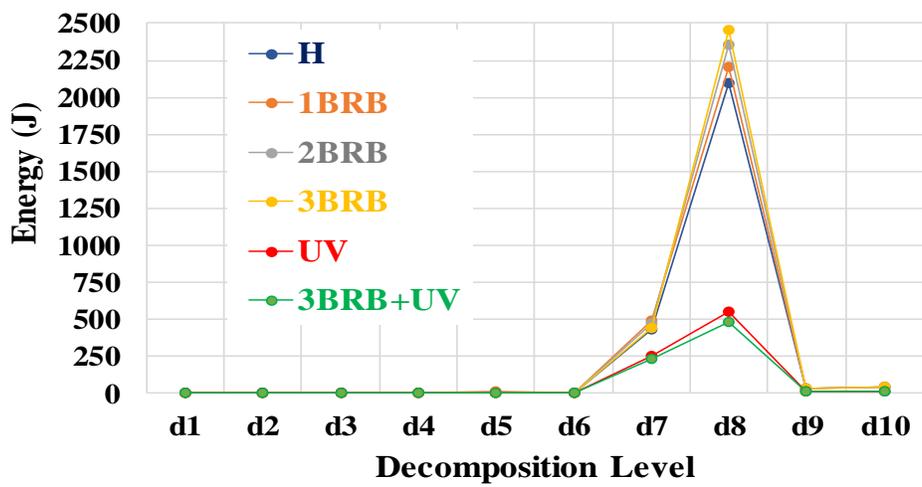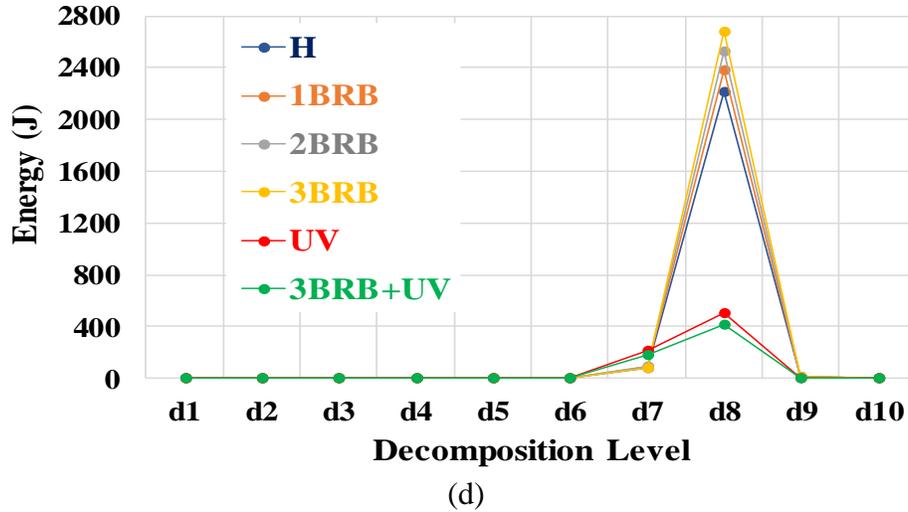
Fig. 4. 7. Energy associated with each decomposition levels for different motor conditions using four different data windows in measured stator current: (a) data window 1; (b) data window 2; (c) data window 3; and (d) data window 4.

Table 4. 2: Threshold and Energy at the decomposition level d8 for all four data windows, each window with 4000 sample points

Data windows		H	1BRB	2BRB	3BRB	UV	3BRB + UV
1	Threshold	4.442	4.638	4.717	4.818	2.287	2.169
	Energy, J	2410.4	2580.8	2720	2870	534.4	470.8
2	Threshold	4.142	4.238	4.417	4.618	2.787	2.645
	Energy, J	1549.5	1647.6	1750.8	1867.9	724.5	640
3	Threshold	4.146	4.24	4.417	4.615	2.781	2.645
	Energy, J	2100.4	2208.3	2350.7	2456.4	545.7	482
4	Threshold	4.118	4.224	4.409	4.513	2.787	2.645
	Energy, J	2210.9	2386.5	2522.4	2682.6	502.4	415.4

4.5 Conclusion

In this paper, stator currents of an 0.25 HP induction motor measured through an experimental test bench under healthy and faulty conditions and 100% loading are analyzed using the DWT for fault diagnosis purpose. Two parameters are evaluated, threshold and energy values, by the DWT processing. It is found

that the threshold value for each decomposition level can serve as a good fault indicator of the motor. Therefore, the results present the effectiveness of the proposed method in the field of motor fault diagnosis application.

References:

[1] 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," *2018 IEEE Industry Applications Society Annual Meeting (IAS)*, 2018.

[2] F. Filippetti, G. Franceschini, and C. Tassoni, "Neural networks aided on-line diagnostics of induction motor rotor faults", *IEEE Trans. Ind. Appl.*, 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. Ind. Appl.*, vol. 31, no. 6, pp. 1280 - 1286, 1995.

[4] 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.

[5] J. CusidÓCusido, L. Romeral, J. A. Ortega, J. A. Rosero, and A. G. A. Espinosa, "Fault Detection in Induction Machines Using Power Spectral Density in Wavelet Decomposition," *IEEE Trans. Ind. Electron.* vol. 55, no. 2, pp. 633–643, 2008.

[6] O. Mohammed, N. Abed, and S. Ganu, "Modeling and Characterization of Induction Motor Internal Faults Using Finite-Element and Discrete Wavelet Transforms," *IEEE Trans. Magn.*, vol. 42, no. 10, pp. 3434–3436, 2006.

[7] N. Bessous, S. E. Zouzou, S. Sbaa, and W. Bentrah, "A comparative study between the MCSA, DWT and the vibration analysis methods to diagnose the dynamic eccentricity fault in induction motors," *2017 6th International Conference on Systems and Control (ICSC)*, 2017.

[8] A. Mohanty and C. Kar, "Fault Detection in a Multistage Gearbox by Demodulation of Motor Current Waveform," *IEEE Trans. Ind. Electron.*, vol. 53, no. 4, pp. 1285–1297, 2006.

[9] 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. Ind. Electron.*, vol. 55, no. 12, pp. 4167–4180, 2008.

[10] A. Glowacz, "Diagnostics of DC and Induction Motors Based on the Analysis of Acoustic Signals," *Measurement Science Review*, vol. 14, no. 5, pp. 257–262, Jan. 2014.

[11] R. Romero-Troncoso, A. Garcia-Perez, D. Morinigo-Sotelo, O. Duque-Perez, R. Osornio-Rios, and M. Ibarra-Manzano, "Rotor unbalance and broken rotor bar detection in inverter-fed induction motors at start-up and steady-state regimes by high-resolution spectral analysis," *Electric Power Systems Research*, vol. 133, pp. 142–148, 2016.

[12] A. Naha, A. K. Samanta, A. Routray, and A. K. Deb, "A Method for Detecting Half-Broken Rotor Bar in Lightly Loaded Induction Motors Using Current," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 7, pp. 1614–1625, 2016.

[13] D. Matić, F. Kulić, M. Pineda-Sánchez, and I. Kamenko, "Support vector machine classifier for diagnosis in electrical machines: Application to broken bar," *Expert Systems with Applications*, vol. 39, no. 10, pp. 8681–8689, 2012.

[14] M. Abd-El-Malek, A. K. Abdelsalam, and O. E. Hassan, "Induction motor broken rotor bar fault location detection through envelope analysis of start-up current using Hilbert transform," *Mechanical Systems and Signal Processing*, vol. 93, pp. 332–350, 2017.

[15] J. Rangel-Magdaleno, H. Peregrina-Barreto, J. Ramirez-Cortes, and I. Cruz-Vega, "Hilbert spectrum analysis of induction motors for the detection of incipient broken rotor bars," *Measurement*, vol. 109, pp. 247–255, 2017.

[16] B. Bessam, A. Menacer, M. Boumehraz, and H. Cherif, "DWT and Hilbert Transform for Broken Rotor Bar Fault Diagnosis in Induction Machine at Low Load," *Energy Procedia*, vol. 74, pp. 1248–1257, 2015.

[17] R. A. Lizarraga-Morales, C. Rodriguez-Donate, E. Cabal-Yepez, M. Lopez-Ramirez, L. M. Ledesma-Carrillo, and E. R. Ferrucho-Alvarez, "Novel FPGA-based Methodology for Early Broken Rotor Bar Detection and Classification Through Homogeneity Estimation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1760–1769, 2017.

[18] J. Pons-Llinares, J. A. Antonino-Daviu, M. Riera-Guasp, M. Pineda-Sanchez, and V. Climente-Alarcon, "Induction Motor Diagnosis Based on a Transient Current Analytic Wavelet Transform via Frequency B-Splines," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1530–1544, 2011.

[19] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to wavelets and wavelet transforms: a primer*. Upper Saddle River: Prentice-Hall, 1998.

[20] 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. Ind. Electron.*, vol. 58, no. 9, pp. 4385–4395, 2011.

[21] "Understanding Wavelets, Part 3: An Example Application of the Discrete Wavelet Transform Video," *Reconstructing an Image from Projection Data - MATLAB & Simulink Example*. [Online]. Available: https://www.mathworks.com/videos/understanding-wavelets-part-3-an-example-application-of-the-discrete-wavelet-transform-121284.html. [Accessed: 08-Jan-2019].

[22] "Wavelet Signal Denoiser," *Reconstructing an Image from Projection Data - MATLAB & Simulink Example*. [Online]. Available: https://www.mathworks.com/help/wavelet/gs/choose-a-wavelet.html. [Accessed: 08-Jan-2019].

Chapter 5

Machine Learning Based Fault Diagnosis for Single- and Multi-Faults for Induction Motors Fed by Variable Frequency Drives

Mohammad Zawad Ali, *Student Member*, IEEE, Md Nasmus Sakib Khan Shabbir, *Student Member*, IEEE, Shafi Md Kawsar Zaman, *Student Member*, IEEE, and Xiaodong Liang, *Senior Member*, IEEE

Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, Newfoundland, Canada.

A version of this chapter has been accepted by 2019 IEEE Industry Application Society (IAS) Annual Meeting. Mohammad Zawad Ali co-authored the paper under the supervision of Dr. Xiaodong Liang. Zawad's contributions in this paper are listed as follows:

- Performed literature searches required for background information of machine learning based fault diagnosis.

- Implemented hardware and performed experiments for two identical induction motors fed by VFDs under single- and multi-faults conditions.

- Implement signal processing DWT technique and machine learning algorithms using experimented data.

- Examined the results and discussed the findings.

- Involved writing the paper as the first author.

Dr. Xiaodong Liang provided the main ideas of the paper, set up the experiment plans, provided continuous technical guidance, checked the results, reviewed the manuscript, provided important suggestions to accomplish the work, and modified the final version of the manuscript. Our group team members, Md Nasmus Sakib Khan Shabbir and Shafi Md Kawsar Zaman, both participated in this research work. Sakib participated in experiments, developed the surface fitting technique and derived feature

calculation formulas, and wrote the relative part of the manuscript. Shafi helped with data processing for feature preparation and machine learning.

In this chapter, the manuscript is presented with altered figure numbers, table numbers and reference formats in order to maintain the thesis formatting guidelines set out by Memorial University of Newfoundland.

Abstract- In this paper, an effective machine learning based fault diagnosis method is developed for induction motors fed by variable frequency drives (VFDs). Two identical 0.25 HP induction motors under healthy, single- and multi-fault conditions were tested in the lab with different VFD output frequencies and motor loadings. The stator current and the vibration of the motors were recorded simultaneously under steady-state for each test, and both signals are evaluated for their suitability for fault diagnosis. The signal processing technique, Discrete Wavelet Transform (DWT), is chosen in this paper to extract features for machine learning. Four families of machine learning algorithms in the MATLAB Classification Learner Toolbox, decision trees, support vector machines (SVM), k-nearest neighbors (KNN), and ensemble, with twenty classifiers are evaluated for their classification accuracy when used for fault diagnosis of induction motors fed by VFDs. To allow fault diagnosis for untested motor operating conditions with different combinations of the motor operating frequency and the motor loading factor, the feature calculation formulas are developed through surface fitting using experimental data for a range of tested frequencies and loadings.

5.1 Introduction

Induction motors are used in various industrial applications due to their reliability, lower cost and ease of control. However, electro-mechanical faults of induction motors can cause severe interruption of industrial processes although protective devices are employed in the system [1]. To overcome such challenges, fault diagnosis approaches for induction motors have been reported in the literature; however, the majority of the research is for induction motors fed directly by the grid [2]-[5].

Recently, due to the advancement of variable frequency drive (VFD) technology and benefits brought by the VFDs, such as flexible production control and soft motor start-up capability, the motor drive systems are increasingly used in various industry facilities [6]. Induction motors driven by VFDs have several differences compared to induction motors directly connected to power sources. These differences are as follows: 1) induction motors fed by VFDs can experience higher stress in bearings and windings because of higher harmonic contents in voltage and current, 2) the operating frequency of the induction motor can be varied at the VFD output, and 3) other factors such as the control method used by the VFD and the carrier frequency of the VFD might also have an effect on induction motor operation. Therefore, it is important to develop an effective fault diagnosis method for induction motors supplied by VFDs.

Very few works have been reported in the literature to investigate fault diagnosis approaches for induction motors fed by VFDs [7]-[27]. The existing research in this particular area can be divided into three categories: 1) model-based approaches [7], 2) signature-extraction-based approaches [8]-[23], and 3) knowledge-based approaches [24]-[27]. In model-based approaches, the mathematical model of induction motor is used to detect and diagnose faults [28]. In signature-extraction-based approaches, signatures extracted from the recorded monitoring signals are used to detect faults. In knowledge-based approaches, machine learning in association with classification learners, and signal-processing techniques are used to detect faults.

Based on our literature review, the only model-based fault diagnosis approach for induction motors driven by VFDs is conducted in [7]. An accurate model of dual-stator winding induction machine (DSWIM) is developed in [7] and the normalized fast Fourier transform (NFFT) of the stator current and control variables is used to investigate the eccentricity impacts. The signature-extraction-based approaches are reported the most. Different signal processing methods have been used to extract signatures including the Fast Fourier Transform (FFT) [8][9][16], the novel time series data mining technique [11], the continuous wavelet transform (CWT)[14], the discrete wavelet transform (DWT) [16], the wavelet packet decomposition (WPD) [17], the diagnostic space vector [18], and the Finite element method [23]. However, the model or signature-extraction-based approaches require a trigger threshold, machine model, and motor or load characteristics, which may not be available or obtained accurately.

On the other hand, the knowledge-based approach uses machine learning to detect faults of induction motors fed by VFDs. It does not need a trigger threshold, machine models, motor or load characteristics; therefore, it is suitable for real-time fault diagnosis once the model is trained. However, only limited research is done in this category for induction motors fed by VFDs [24]-[27]. In [24][25], the advanced complex wavelets transform is used for feature extraction, M-SVM and k-nearest neighbors (KNN) are used for multiple fault detection and isolation for a VFD driven induction motor. An experimental comparative evaluation of different machine learning techniques is carried out in [26]. Classification accuracy among six machine learning algorithms, namely Bayesian Learning, Instance-Based Methods, Bootstrap Aggregating, Boosting Algorithms, Artificial Neural Networks, Support Vector Machines (SVM), are compared in the analysis. Although the artificial neural network (ANN) is very popular in fault diagnosis for induction motors fed directly by power grid [28], only Ref [27] is found in the literature using ANN in fault diagnosis for induction motors fed by VFDs. Short time Discrete Fourier transform (STDFT) is used in [27] to extract features which are used to train ANN.

There are three major challenges related to fault diagnosis for induction motors fed by VFDs: 1) What signal should be chosen for signal processing? 2) Can we deal with single- and multi-faults? 3) How the prediction can be effectively made for situations that the testing data are not available for training?

Challenge 1: The stator current of the motor, either alone or combined with other parameters, are commonly used for signal processing for signature-extraction or knowledge-based approaches. To extract features, stator current alone is used in [8][12][14][17][22][26], the combined stator voltage and current are used in [7][27], the stator current and estimated mechanical speed are both used in [10]. In addition to the stator current, the machine vibration is used in [24], the instantaneous input power is used in [18][21]. Although the stator current and vibration signals are commonly used for fault diagnosis of induction motors fed directly by power grids, no comparative analysis has been carried out for induction motors fed by VFDs.

Challenge 2: In the literature, induction motor fault diagnosis is reported mostly for single fault, such as the eccentricity [7][14][27][21], bearing fault [17][10], rotor faults [15][18][26][22][23], stator winding fault [8][9][12], broken rotor bar/end-ring and eccentricity in [16], broken rotor bars (BRBs) and broken end-ring connectors [11], and stator winding and bearing faults [24]. None of these investigations consider the impact of the multiple faults' occurrence at the same time.

Challenge 3: In real life, the loading factor and the operating frequency of induction motors can be different from the values used in testing. No guidelines are proposed to determine features for untested conditions in order to train machine learning algorithms for induction motors fed by VFDs.

In this paper, to address these challenges, a robust machine learning based fault diagnosis method is proposed for a wide variety of single- and multi-faults of VFD driven induction motors. The stator current and vibration signals recorded simultaneously in the lab for a wide range of operating frequencies and load factors of the induction motor are processed using the DWT to extract features for machine learning. The

major contributions of the paper are summarized as follows: 1) Compare the fault diagnosis performance using the measured stator current and vibration signals; 2) Compare the fault classification performance for four families of machine learning algorithms (Decision trees, SVM, KNN, and Ensemble in MATLAB Classification Learner toolbox); 3) Create several single- multi-faults for the tested motors to evaluate the robustness of the proposed method; 4) Determine and validate features for the untested cases for training, the equations to calculate these unknown features are developed using surface fitting through MATLAB curve fitting tool box.

The paper is arranged as follows: the proposed machine learning based fault diagnosis approach using experimental data for induction motors fed by VFDs is given in Section 5.2; detailed experimental set-up is provided in Section 5.3; in Section 5.4, signal processing of the measured stator current and vibration signals using the DWT is conducted, and eight features are extracted through DWT processing; details about the machine learning classifiers are provided in Section 5.5; classification accuracies using different classifiers are demonstrated in Section 5.6; in Section 5.7, the surface fitting equations are developed to calculate unknown features vs. motor loadings and operating frequencies; and conclusions are drawn in Section 5.8.

5.2 The proposed Fault Diagnosis Approach

In this paper, an effective fault diagnosis approach for VFD supplied induction motors using experimental data is proposed in Fig. 5.1. The basic procedure is configured under three considerations: 1) Hardware implementation; 2) Simulation-based implementation and analysis; and 3) Quantitative comparison and decision.
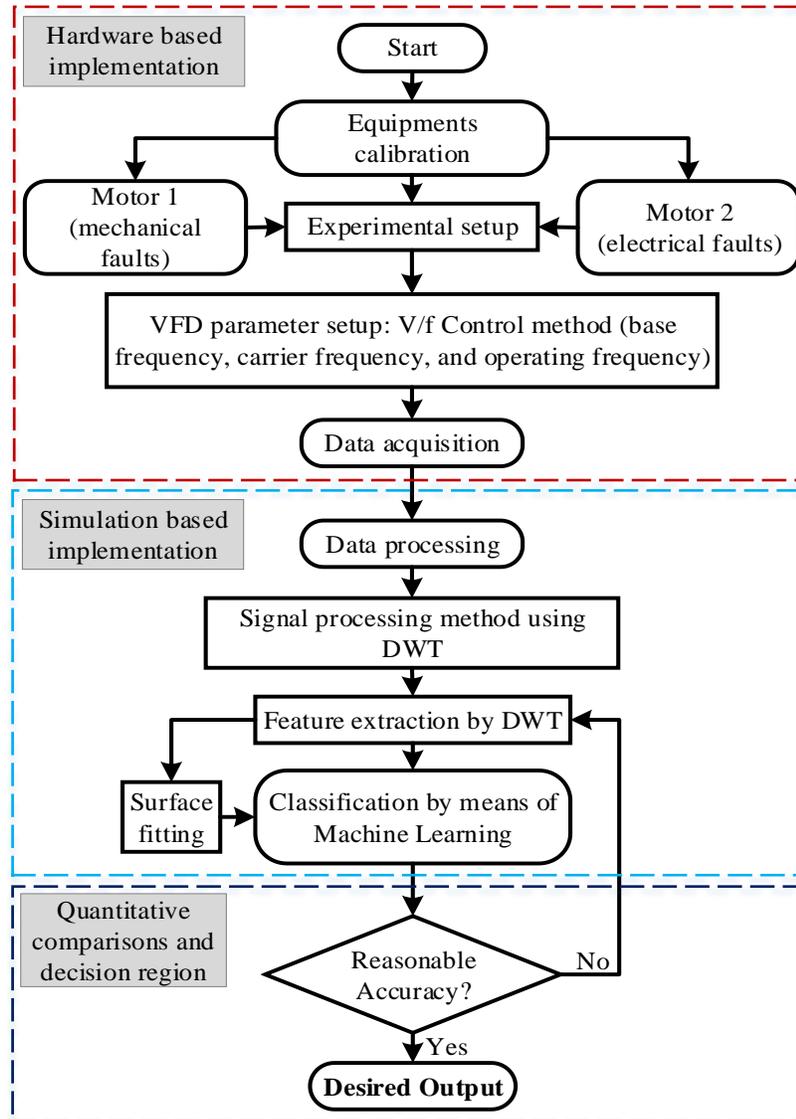
Fig. 5. 1. The flow chart of the proposed method.

The approach can be implemented in a six-step procedure: 1) conduct experiments for an induction motor fed by a VFD under healthy, single- and multi-fault conditions by considering different output frequencies of the VFD and load factors of the motor; 2) record stator currents and vibration signals simultaneously using a power quality analyzer and vibration sensors; 3) choose a suitable signal processing method for features extraction, such as DWT; 4) compare different feature selections and determine the most suitable features for the system; 5) classify faults for the motor using the chosen classifiers; and 6) develop surface

fitting equations to calculate features vs. motor loadings and operating frequencies for conditions without measurement data.

5.3 Experimental Set-Up

In the experiment set-up, two identical 4-pole, 0.25 HP, 208-230/460V, 1725 rpm rated squirrel-cage induction motors (Model LEESON 101649) are tested fed by a VFD in the lab. The motors are named as "Motor 1" and "Motor 2" and are treated as sister units. The VFD is manufactured by Saftronics (Model: CIMR-G5U23P7F). The input ratings include AC three phase, 200-220 V at 50 Hz (200-230V at 60Hz), and 21 A. The output ratings include AC three phase, 0-230 V, 0-400 Hz, and 17.5 A.

As shown in Fig. 5.2, the faults applied on Motor 1 are mainly mechanical faults, and the faults on Motor 2 are electrical faults. The Motor 1 testing include: (a) a healthy condition (H), (b) an unbalance shaft rotation (UNB), (c) a bearing fault (BF), (d) the combined BF and UNB fault, (e) the combined BF and one broken rotor bar (BRB) faults, and (f) the combined BF, UNB, and unbalance voltage (UV) condition from the three-phase power supply. The Motor 2 testing include: (a) a healthy condition (H), (b) a UV from three-phase power supply, (c) one BRB fault, (d) two BRBs fault, (e) three BRBs fault, and (f) the combined UV and three BRB fault.
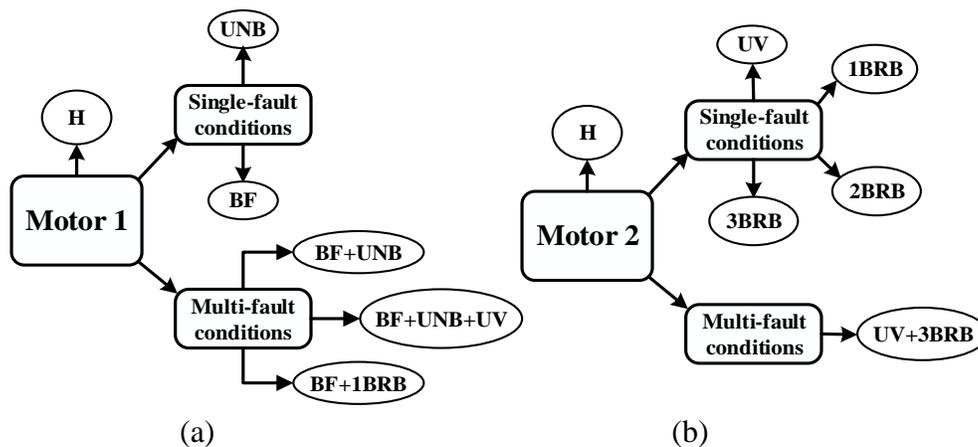
Fig. 5. 2. The testing for the two motors: (a) Motor 1; (b) Motor 2.

The experimental test bench is shown in Fig. 5.3. The induction motor is connected through a VFD to a three-phase power supply. The load is a dynamometer coupled to the motor shaft through a belt pulley. Motor loadings can be adjusted by the dynamometer's control knob. Under the full load, the torque of the motors is 7 pound force inch (lbf-in) at the rated speed.
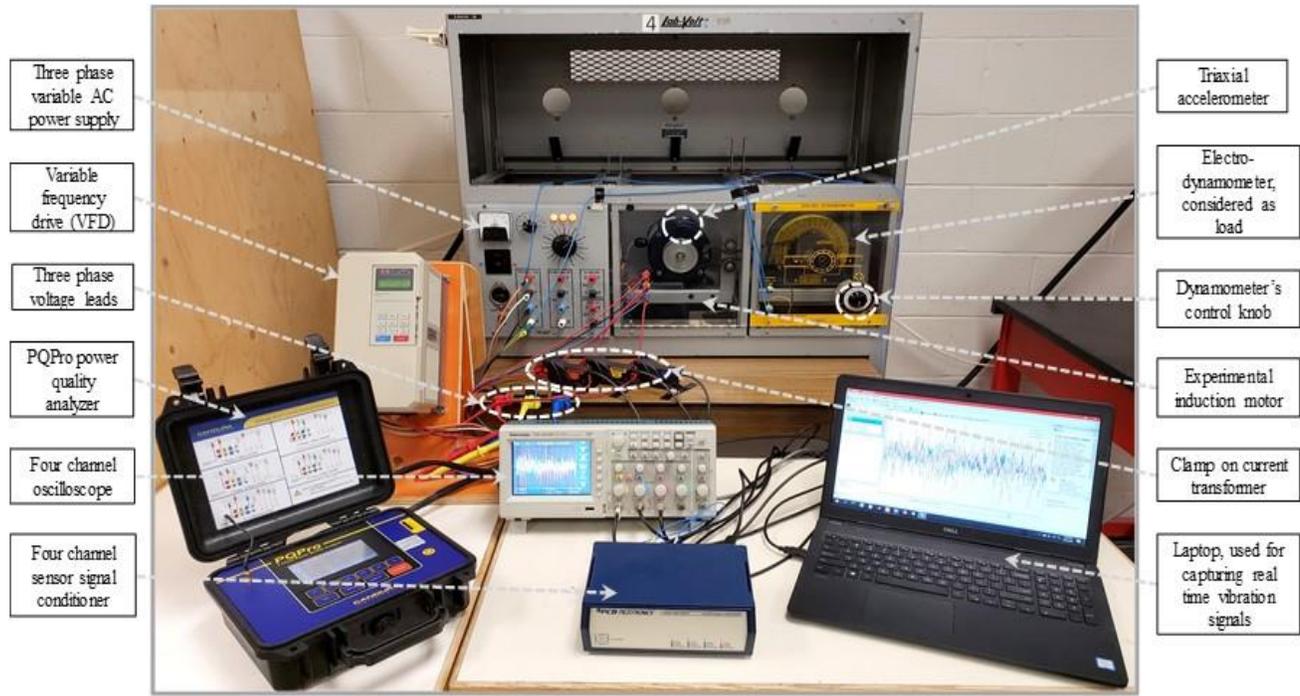
Fig. 5. 3. The experimental test bench for induction motors fed by a VFD.

Fig. 5.4 shows the schematic diagram of the system set-up. An eight-channel power quality analyzer, PQPro by CANDURA instrument, is used to measure the three-phase stator currents (I1, I2, and I3) of the motor. The measurements are taken on the output of the VFD. A tri-axial accelerometer (Model 356A32) with a four-channel sensor signal conditioner (Model 482C05) mounted on top of the motor near the face end is used to record vibration signals.

It is specified in this measurement that the vibration at the axial direction is x-axis, at the vertical direction is y- axis, and at the horizontal direction is z-axis. A four-channel oscilloscope is patched between the sensor signal conditioner and the computer for vibration data acquisition. The sampling frequency for

vibration measurements is 1.5 kHz. The stator currents and vibration signals were measured simultaneously under steady-state operating conditions.
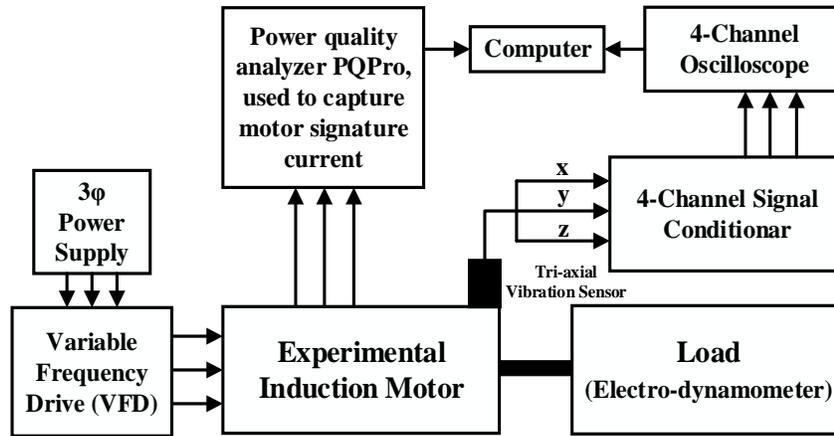
Fig. 5. 4. Schematic diagram of the system set-up.

Photos of the motors with a specific type of fault used in the experiment are shown in Fig. 5.5. A BRB fault was realized by drilling a hole of a 5 mm diameter and 18 mm depth in the rotor bar (Fig. 5(a)). Two and three holes were drilled on adjacent rotor bars for two BRBs and three BRBs faults, respectively (Figs. 5.5 (b) and 5.5 (c)). The general roughness type of bearing fault was realized by a sand blasting process, the outer and inner raceway of the bearing becomes very rough (Fig. 5.5 (d)). The UNB was created by adding extra weight on part of the pulley (Fig. 5.5 (e)). An UV condition was formed by adding an extra resistance on the second phase of the VFD output.

(a)   (b)   (c)

<center>(d)   (e)</center>
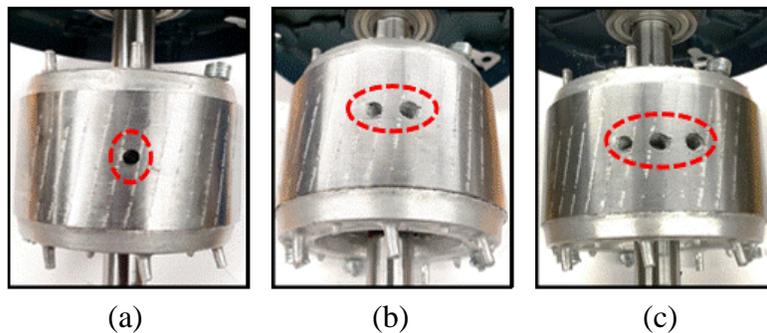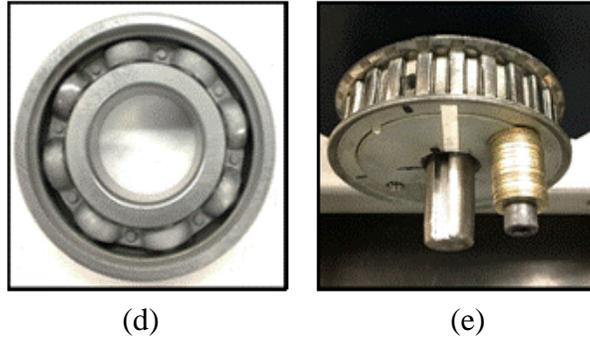
Fig. 5. 5. Photos of faults applied on the motors in experiments: (a) 1 BRB, (b) 2 BRBs, (c) 3 BRBs, (d) the general roughness type of bearing fault, and (e) the UNB condition.

Six output frequencies of the VFD from 45 Hz to 70 Hz were used in the testing. Four different carrier frequencies (1100 Hz, 3100 Hz, 8000 Hz, and 15 kHz) were evaluated, and 3100 Hz was chosen as the carrier frequency for all testing. Six different loadings ranging from no load (0%) to full load (100%) of the motors were tested for each output frequency per fault. Table 5.1 summarizes various parameters used in the experiments. By combining different types of faults, output frequencies of the VFD and motor loadings, a total of 540 tests were conducted in the lab.

<center>Table 5. 1: The equipment settings of the experiments</center>

Parameters	Settings
VFD output frequency, Hz	45, 50, 55, 60, 65, 70
VFD carrier frequency, Hz	3100
VFD base frequency	(1) 60 Hz for output frequency at or below 60 Hz (2) 65 Hz for 65 Hz output frequency (3) 70 Hz for 70 Hz output frequency
VFD control method	Voltage per Hz control
Motor loading factor, %	0, 20, 40, 60, 80, 100

5.4 Signal Processing Using DWT for Feature Extraction

The wavelet transform is an effective way to define a signal that is comprised of different frequency components by decomposing a signal into wavelets, which are confined by time and frequency. The discrete wavelet transform (DWT) can be used to analyze a non-stationary signal in time-frequency domain

[29][30]. The DWT uses orthogonal wavelets like Daubechies wavelet series for decomposing the signal into several frequency bands. Because of this feature, it is also known as multiresolution analysis [31].

In this paper, the DWT method is adopted for feature extraction through MATLAB Wavelet toolbox. Among all other different wavelet families in the DWT analysis, the wavelet from Daubechies family with four vanishing moment as db4 is considered as the mother wavelet with the 6th level decomposition. Eight statistical features (mean, median, standard deviation, median absolute deviation, mean absolute deviation, L1 norm, L2 norm, and the maximum norm) are evaluated for the motor stator currents and vibration signals processed by DWT. These features are tabulated in Table 5.2 [32][33], which will be used for machine learning.

Table 5.3 shows a sample of features obtained using the z-axis vibration signal for Motor 1 with a bearing fault (100% loading factor of the motor and 60 Hz drive output frequency). Every set of eight features, such as S1 in the first row of Table 5.3, is obtained by choosing a data window, which contains 9000 sample points, from the z-axis vibration signal and processed by the DWT. Other nine sets of features (from S2 to S10) are determined by taking sample points in a similar way from nine different data windows. Similarly, Table 5.4 shows a sample of features obtained using the stator current I2 for Motor 2 with a 1 BRB fault (80% loading factor of the motor and 50 Hz drive output frequency).

Fig. 5.6 shows one feature, Mean, for Motors 1 and 2 processed by the current I2 vs. motor loading factors and different types of faults for a fixed output frequency of 60 Hz. Other features show similar patterns to Fig. 5.6. Fig. 5.7 shows one feature, Mean, for Motors 1 and 2 processed by the z-axis vibration signal vs. VFD output frequencies and different types of faults for a fixed motor loading factor of 60%. Other features show similar patterns to Fig. 5.7.

Fig. 5.8 shows the processed one phase stator current signal I2 for Motor 2 under a 1 BRB fault (100% motor loading and 60 Hz drive output frequency). Fig. 5.9 shows the processed z-axis vibration signal for Motor 2 under a 1 BRB fault (40% motor loading and 50 Hz drive output frequency).

Table 5. 2: Potential Statistical features [28] [32] [33]

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	$Median\_AD = median(\|x_i - meidan(X)\|)$
Mean Absolute Deviation	$Mean\_AD = \frac{1}{N}\sum_{i=1}^{N} \|x_i - \mu_x\|$
L1 norm	$\|L\|_1 = \sum_{i=1}^{N} \|x_i\|$, the sum of absolute values of its components, also known as one-norm, or mean norm
L2 norm	$\|L\|_2 = \sqrt{\sum_{i=1}^{N} \|x_i\|^2}$, 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)	$\|L\|_\infty = \max\{\|x_i\|: i = 1, 2, ..., n\}$, the maximum of absolute values of its components, also known as infinity norm, or uniform norm.

Table 5. 3: Potential Features using Z-axis vibration signal (Motor 1, BF, 100% loading, 60 Hz drive output frequency)

Features	Mean	Median	Std. Dev.	Median Absolute Dev.	Mean Absolute Dev.	L1 norm	L2 norm	Max norm
s1	0.003018	0	0.07637	0.04	0.05462	541.00	7.250	1.28
s2	0.020240	0.04	0.08115	0.04	0.05763	527.00	7.934	1.76
s3	0.034740	0.04	0.08405	0.04	0.05260	539.70	8.628	2.16
s4	0.019210	0.04	0.07822	0.04	0.05709	517.80	7.640	1.68
s5	0.027120	0.04	0.07440	0.04	0.05391	517.90	7.513	1.48
s6	0.024430	0.04	0.07383	0.04	0.05366	531.30	7.377	1.92
s7	0.006244	0	0.08218	0.04	0.05701	511.10	7.818	1.72
s8	0.003649	0	0.08302	0.04	0.05654	504.00	7.883	1.52
s9	0.010920	0	0.07050	0.04	0.05119	519.60	6.768	1.16
s10	0.004240	0	0.07388	0.04	0.05283	526.40	7.020	1.92

Table 5. 4: Potential features using stator current I2 (Motor 2, 1 BRB, 80% loading, 50 Hz drive output frequency)

Features	Mean	Median	Std. Dev.	Median Absolute Dev.	Mean Absolute Dev.	L1 norm	L2 norm	Max norm
s1	0.004439	0.006474	0.8838	0.95920	0.7754	6979	83.840	1.526
s2	0.004178	0.009441	0.8964	0.96970	0.7892	7103	85.040	1.533
s3	0.006279	0.007823	0.8839	0.96570	0.7769	6992	83.850	1.553
s4	0.004808	0.005665	0.8932	0.96460	0.7839	7055	84.730	1.530
s5	0.005496	0.007823	0.8900	0.96950	0.7832	7049	84.430	1.527
s6	0.004785	0.007013	0.8894	0.95440	0.7793	7014	84.370	1.528
s7	0.006381	0.011870	0.8914	0.97000	0.7852	7067	84.560	1.516
s8	0.006344	0.000540	0.8872	0.95890	0.7771	6994	84.160	1.524
s9	0.005027	0.008632	0.8951	0.96890	0.7880	7092	84.920	1.525
s10	0.005090	0.007013	0.8818	0.95760	0.7726	6954	83.650	1.507
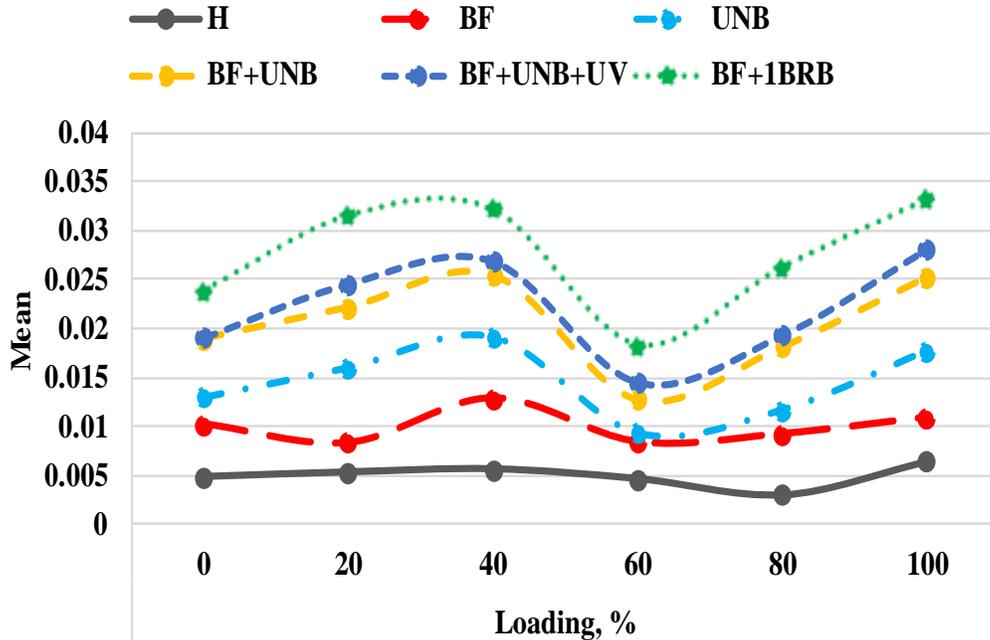
(a)
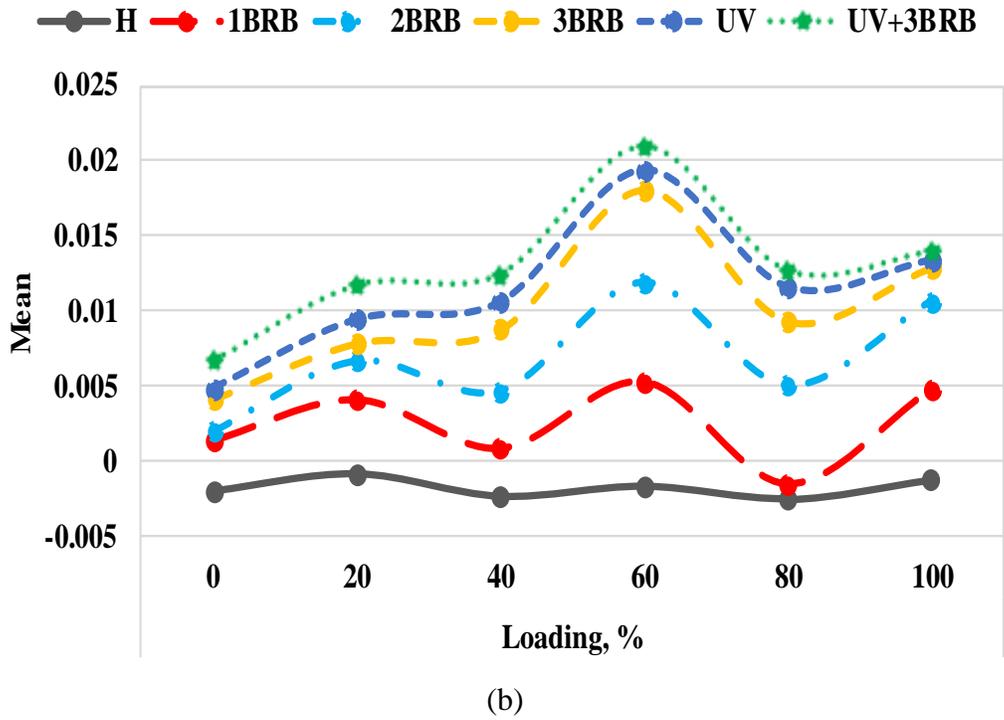
Fig. 5. 6. One feature, Mean, vs. motor loadings and different types of faults using one phase stator current signal I2 (60Hz output frequency): (a) Motor 1; (b) Motor 2.

(b)
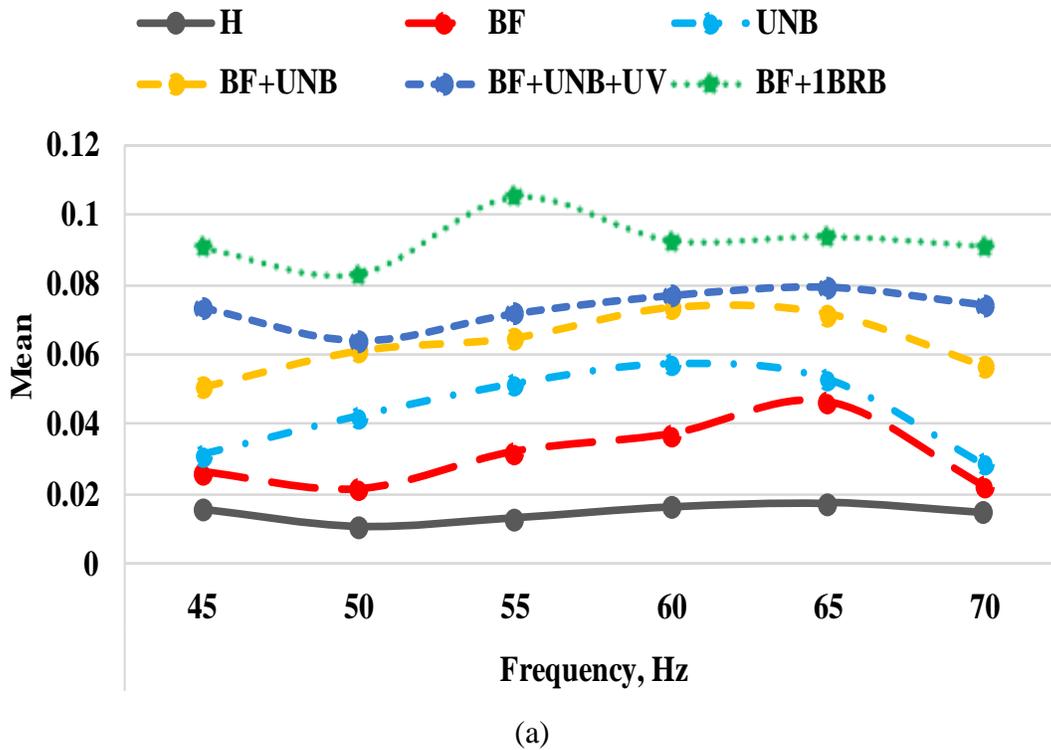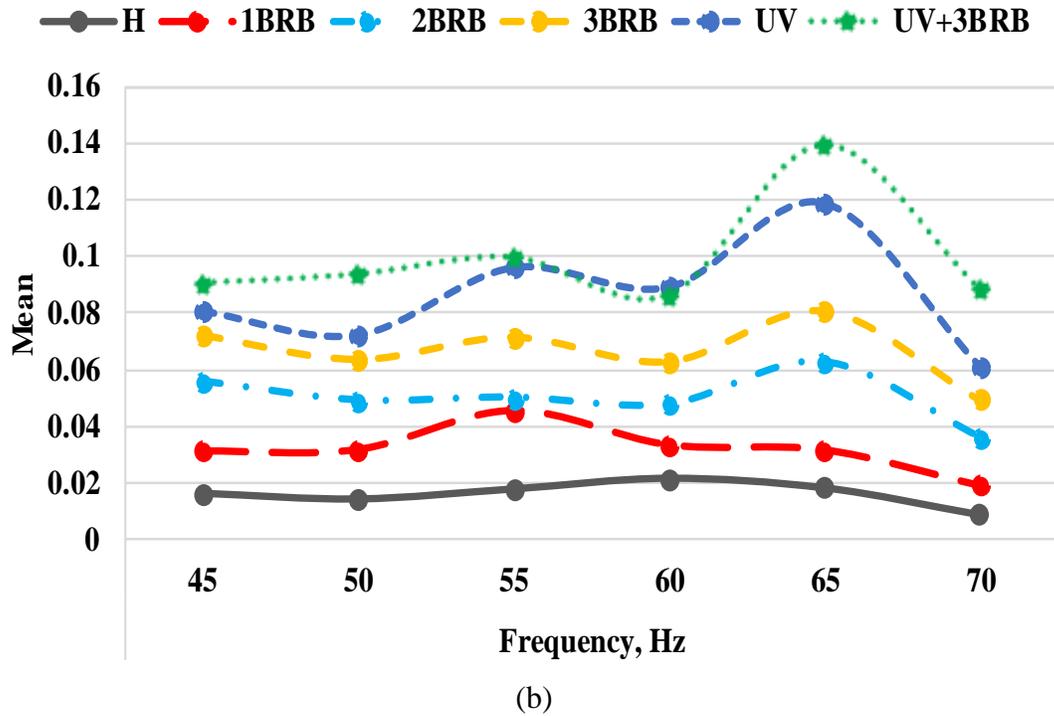
Fig. 5. 7. One feature, Mean, vs. VFD output frequency and different types of faults using z-axis vibration signal (60% motor loading): (a) Motor 1; (b) Motor 2.
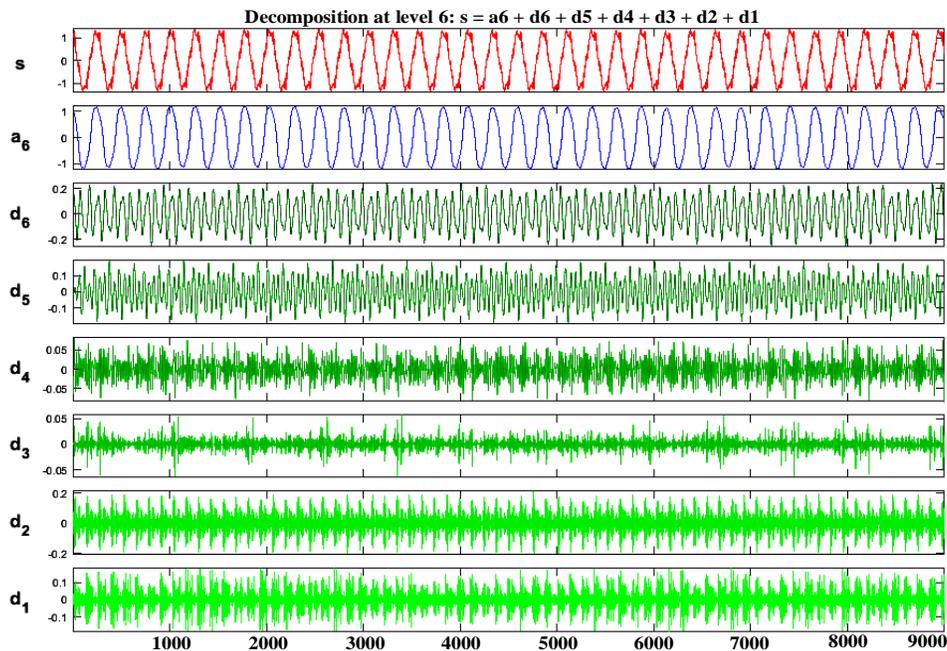
Fig. 5. 8. The processed one phase stator current signal I2 using DWT for Motor 2 under a 1 BRB fault (100% motor loading and 60 Hz drive output frequency).
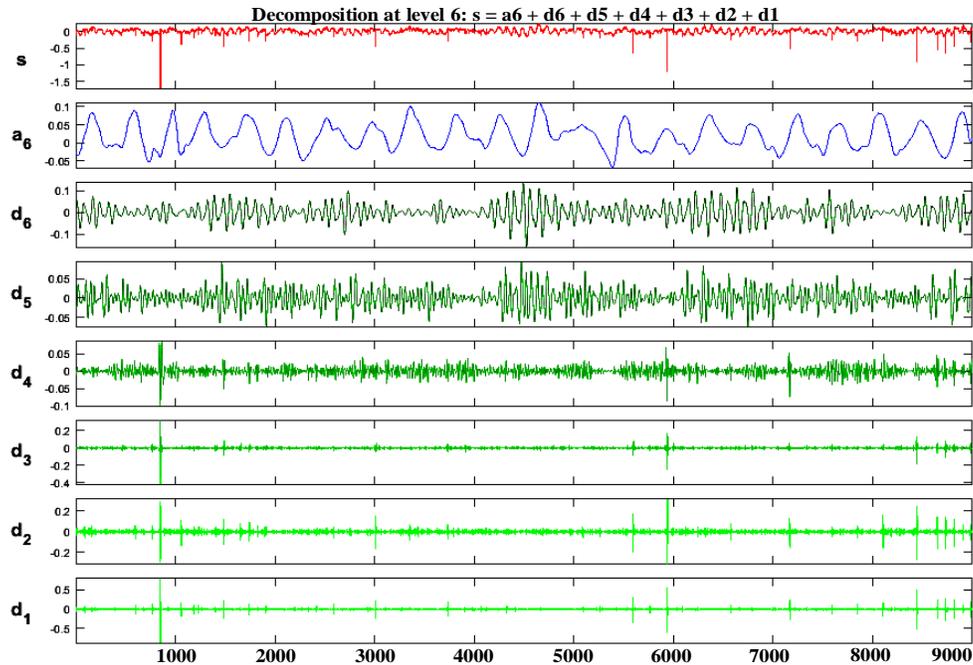
Fig. 5. 9. The processed z-axis vibration signal using DWT for Motor 2 under a 1 BRB fault (40% motor loading and 50 Hz drive output frequency).

5.5 Machine Learning Classifiers

In this paper, four families of classification algorithms offered in the MATLAB Classification Learner Toolbox including Decision Trees, SVM, KNN, ensemble are selected to evaluate their suitability for fault diagnosis of induction motors fed by VFDs, where twenty different classifiers are chosen for evaluation.

5.5.1 Classification Algorithms

The decision tree learning is a classification method by using a decision model to predict and evaluate possible consequences and event outcome. The algorithms hold conditional control statements and are used as descriptive means for calculating conditional probabilities. A decision tree mainly consists of three nodes: decision nodes, chance nodes, and end nodes. Decision nodes represent the root the model, chance nodes represent the possible event outcomes, and end nodes provide the classification [34] [35].

SVM is a commonly used machine learning based data classification and regression tool, which generally classifies a dataset into two classes; positive and negative classes. The two classes are separated by hyperplane. Kernel functions in SVM are used for nonlinear transformation [36]-[40]. The common types of kernel functions like linear kernel, polynomial kernel, Gaussian or radial basis function (RBF) kernel, are used in this study.

KNN is an instance-based classification technique, where the learner summarizes the training data but does not abstract any information from the training data. The basic classification pattern follows by an unknown instance by correlating with a known instance via a valid distance or a similarity functions. In KNN, input data set is separated into a fixed number (k) of clusters and the center of the cluster is called centroid. A centroid is a data point that can be either real or imaginary. All centroids are used to train the KNN classifier. The emanated classifier is proposed during the initialization of primary cluster. The process between classification and centroid adjustment is repeated until the value of centroid become steady and later, these stabilized centroids are used for the clustering of input data. Therefore, the transformation of an anonymous dataset into a known one is acquired by stabilized centroids [37][41].

An ensemble is a superior classifier and uses multiple algorithms to enhance its performance and prediction accuracy. It combines multiple diverse single classifiers. In ensemble classifier, the trained ensemble represents a single hypothesis. This hypothesis does not necessarily need to be presented within the set of hypothesis space from where it is initiated. Due to this flexibility, sometimes it tends to over-fit the training data. Some ensemble methods like Bagged Trees tend to reduce over fitting of training data. The weaker error correlation between single classifiers gives better prediction accuracy [42]-[45].

5.5.2 Classifiers from the Toolbox

In this paper, the following four families of classification algorithms in MATLAB Classification Learner toolbox are chosen to perform fault diagnosis with twenty classifiers:

- Decision trees: Complex Tree, medium Tree, and simple Tree.

- 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 5.5 shows the description of each classifier used in the paper. A five-fold cross validation for all classifiers is performed to prevent the model from overfitting.

Table 5. 5: Description of twenty Classifiers in MATLAB Classification Learner Toolbox

Classification algorithms	Classifier	Classifier description from MATLAB classification learner toolbox
Decision Trees	Fine Tree	A decision tree with many leaves that make many fine distinctions between classes, where maximum number of splits is 100.
	Medium Tree	A decision tree of medium flexibility with fewer leaves, where maximum number of splits is 20.
	Coarse Tree	A simple decision tree with few leaves that makes coarse distinctions between classes, where maximum number of splits is 4.
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.
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	Boosted trees	This model creates an ensemble of medium decision trees using the AdaBoost algorithm. Compared to 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.

5.6 Classification Results for Various Faults

5.6.1 Fault Diagnosis Results

Fed by a VFD, the motor operates at a frequency between 45 Hz and 70 Hz, and operates under different loading conditions between 0 % and 100% in the lab testing. It is important to evaluate the accuracy of fault diagnosis of induction motors with such variations. Fig. 5.10 shows the fault classification accuracy for all faults implemented on Motor 1 at 100% motor loading and 60 Hz operating frequency using the stator current I2 and z-axis vibration signals.

Similarly, Fig. 5.11 presents the fault diagnosis accuracy for all faults implemented on Motor 2 at 80% loading and 45 Hz operating frequency using the stator current I2 and z-axis vibration signals. The corresponding accuracy data for Fig. 5.10 and Fig. 5.11 are provided in Tables 5.6 and 5.7.

It is found that the stator current demonstrates a significantly better performance than the vibration for both motors. The accuracy values using vibration signal are mostly below 70% for Motor 1, and 60% for Motor 2; while the accuracy values using the stator current signal can be as high as 92.8% for Motor 1, and 100% for Motor 2. In real life applications, the stator current is much easier to measure than the vibration signal. Therefore, the stator current signal is recommended to be used for induction motor fault diagnosis fed by VFDs.

Motor 2 with electrical faults has much better accuracy than Motor 1 with mechanical faults when both using the stator current. Among the 20 classifiers, 3 classifiers (Linear SVM, Medium Gaussian SVM, and Subspace Discriminant) has above 90% accuracy for Motor 1, and 8 classifiers (Fine Tree, Medium Tree, Linear SVM, Quadratic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Bagged Trees, and Subspace KNN) has above 90% accuracy for Motor 2.
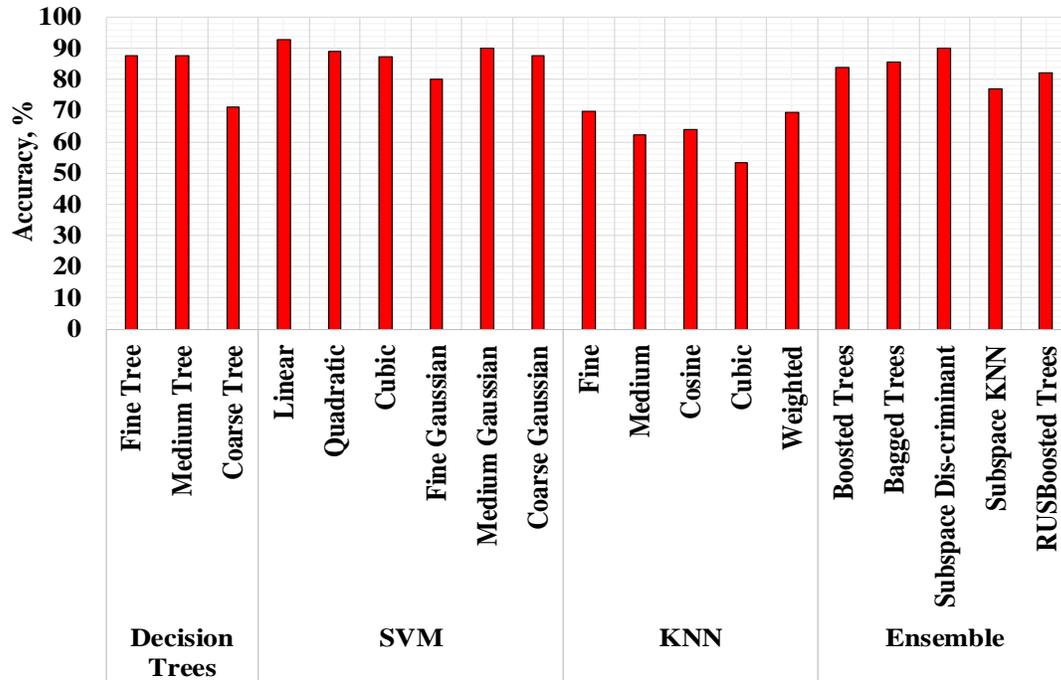
(a)
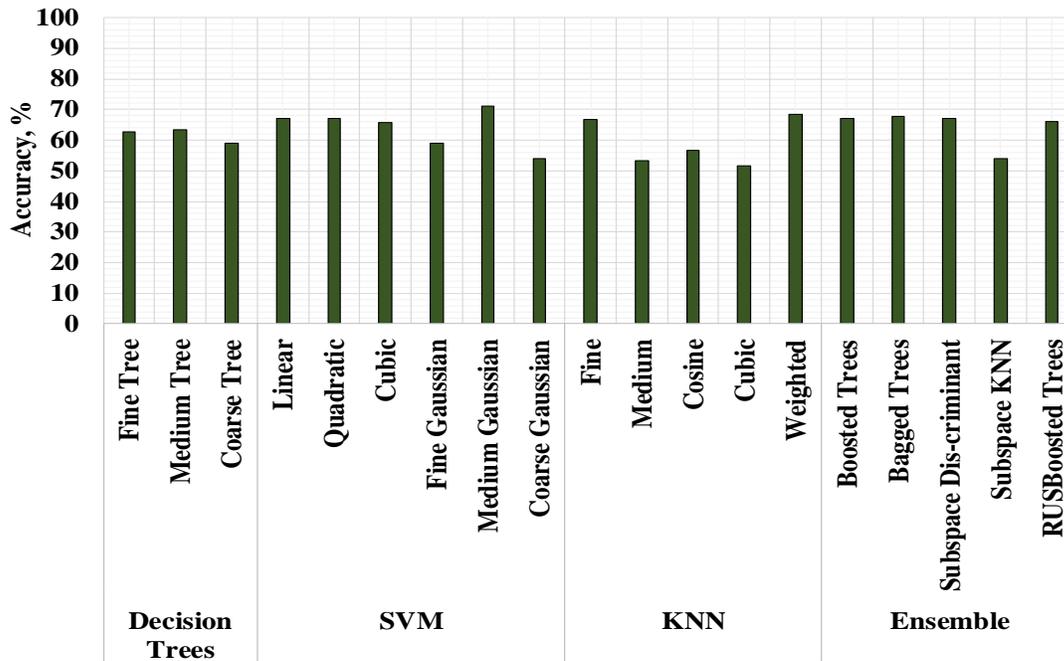
(b)

Fig. 5. 10. Classification accuracy for all faults implemented on Motor 1 (100% loading and 60Hz): (a) stator current I2; (b) z-axis vibration.
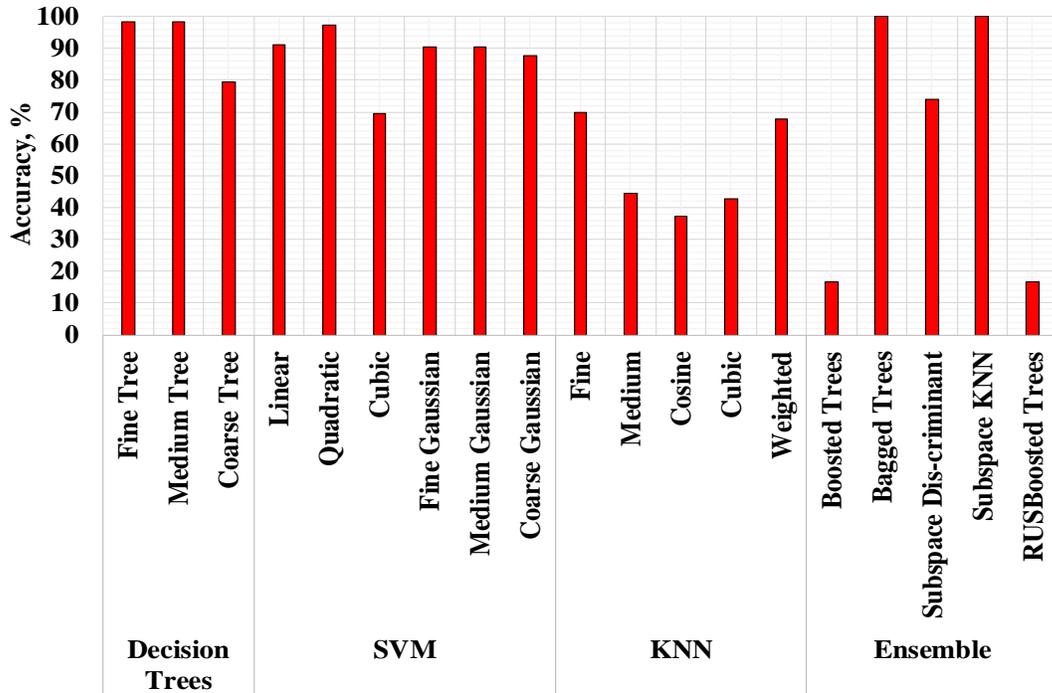
(a)
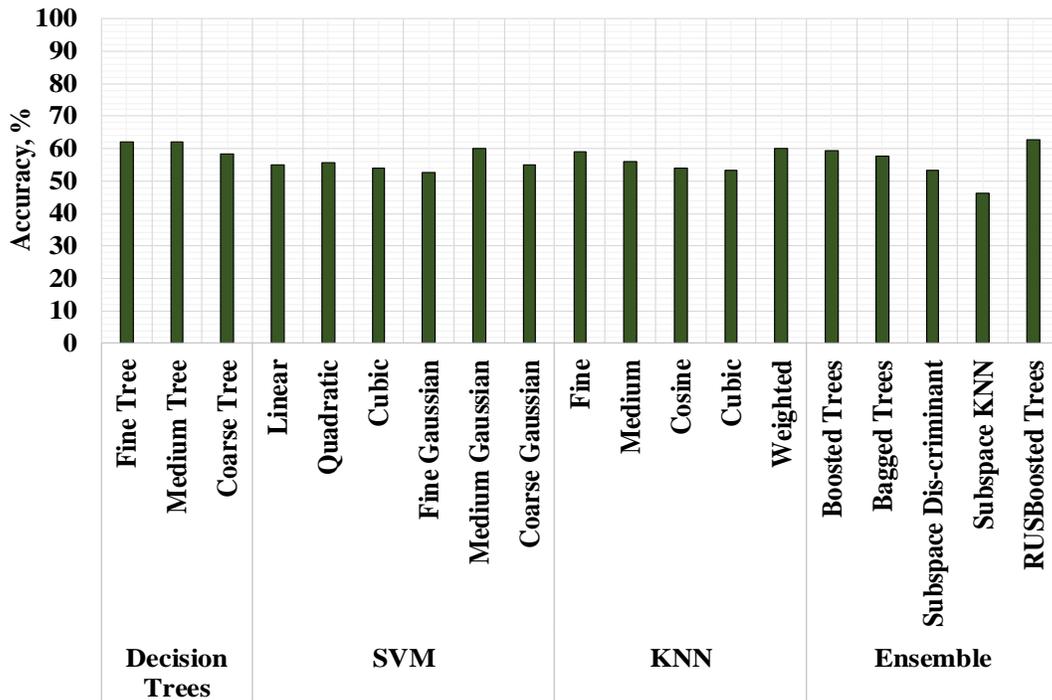
(b)

Fig. 5. 11. Classification accuracy for all faults implemented on Motor 2 (80% loading and 45Hz): (a) stator current I2; (b) z-axis vibration.

Table 5. 6: Accuracy for classification of all faults for Motor 1 (100% loading and 60Hz) using various classifiers

Classification Method	Classifiers	Classification accuracy, %	
		Current (I$_2$)	Vibration (z-axis)
Decision Trees	Fine Tree	87.8	62.8
	Medium Tree	87.8	63.3
	Coarse Tree	71.1	58.9
SVM	Linear SVM	**92.8**	67.2
	Quadratic SVM	88.9	67.2
	Cubic SVM	87.2	65.6
	Fine Gaussian SVM	80	58.9
	Medium Gaussian SVM	**90**	71.1
	Coarse Gaussian SVM	87.8	53.9
KNN	Fine KNN	70	66.7
	Medium KNN	62.2	53.3
	Cosine KNN	63.9	56.7
	Cubic KNN	53.3	51.7
	Weighted KNN	69.4	68.3
Ensemble	Boosted Trees	83.9	67.2
	Bagged Trees	85.6	67.8
	Subspace Discriminant	**90**	67.2
	Subspace KNN	77.2	53.9
	RUSBoosted Trees	82.2	66.1

Table 5. 7: Accuracy for classification of all faults for Motor 2 (80% loading and 45Hz) using various classifiers

Classification Method	Classifiers	Classification accuracy, %	
		Current (I2)	Vibration (z-axis)
Decision Trees	Fine Tree	**98.3**	62.2
	Medium Tree	**98.3**	62.2
	Coarse Tree	79.4	58.3
SVM	Linear SVM	**91.1**	55
	Quadratic SVM	**97.2**	55.6
	Cubic SVM	69.4	53.9
	Fine Gaussian SVM	**90.6**	52.8
	Medium Gaussian SVM	**90.6**	60
	Coarse Gaussian SVM	87.8	55
KNN	Fine KNN	70	58.9
	Medium KNN	44.4	56.1
	Cosine KNN	37.2	53.9
	Cubic KNN	42.8	53.3

	Weighted KNN	67.8	60
	Boosted Trees	16.7	59.4
	Bagged Trees	**100**	57.8
Ensemble	Subspace Discriminant	73.9	53.3
	Subspace KNN	**100**	46.1
	RUSBoosted Trees	16.7	62.8

The classifier performance is evaluated by employing the confusion matrix and receiver operating characteristic (ROC) curve. The confusion matrix is able to recognize the regions, where the classifier has performed correctly or poorly, and to evaluate how a classifier is executed in each class. In this study, the confusion matrix has been summarized by choosing the positive predictive value (PPV) and false discovery rate (FDR). The PPV is shown in green for the correctly predicted points in each class, and the FDR is shown below the PPV in red for incorrectly predicted points in each class. The accuracy from confusion matrix is calculated by

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{1}$$

Where, TP is true positive, TN is true negative, FP is false positive, and FN is false negative. The ROC curve condenses the overall performance of a classifier over all possible threshold. The area under the curve (AUC) gives a brief perception about how confidently the classification is done. The ROC curve plots the true positive rate (TPR) as a function of the false positive rate (FPR). The ROC is a graphical representation of confusion matrix where parameters are calculated as follows [46] [47]:

$$TPR = \frac{TP}{TP+FN} = 1 - false\ negative\ rate \tag{2}$$

$$FPR = \frac{FP}{FP+TN} = 1 - true\ negative\ rate \tag{3}$$

The TPR implies how often the classifier predicts positive when the actual classification is positive, while the FPR signifies how often the classifier incorrectly predict positive when the actual classification is negative. Both TPR and FPR are ranges from 0 to 1 and AUC ranges from 0.5 to 1. An AUC value of 1

denote a better result with no misclassified points and 0.5 represents that the classifier is no better than random estimation. Fig. 5.12 illustrates the confusion matrix and ROC curve with the 100% classification accuracy achieved by the classifier, Subspace KNN, for Motor 2 at 80% motor loading and 45 Hz operating frequency using the stator current I_2.

(a)

(b)
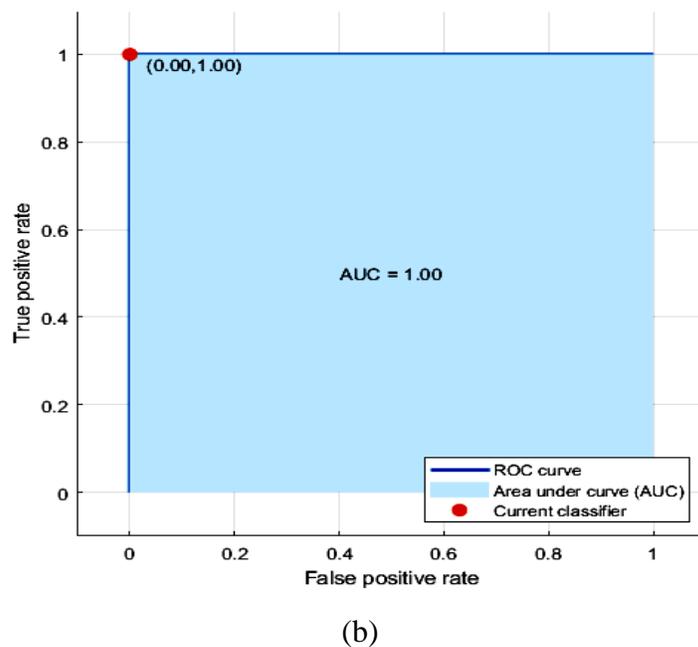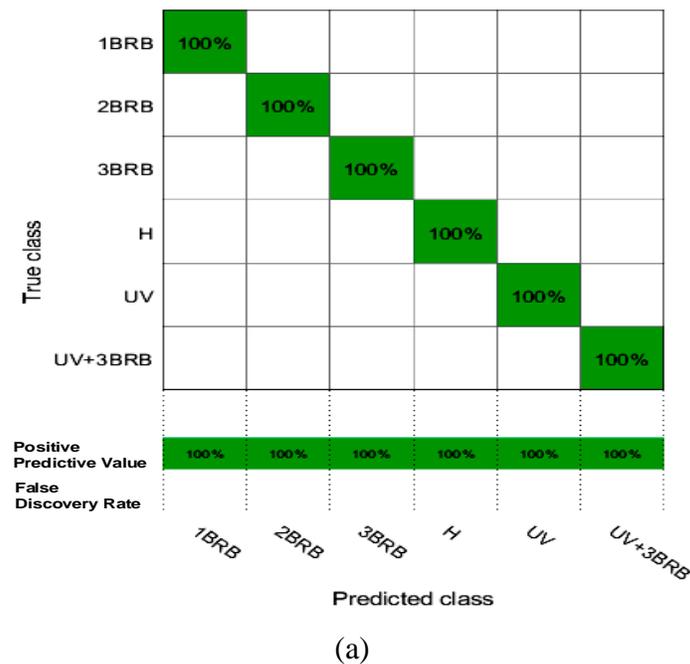
Fig. 5. 12. Classification accuracy using Subspace KNN for Motor 2 at 80% motor loading and 45 Hz using the current I2: (a) confusion matrix; (b) ROC curve.

5.6.2 Influence of the Number of Chosen Features

In this study, eight statistical features are considered as features for fault classifications. It is essential to examine the influence of the number of features on the classification accuracy. Six cases are considered as follows: Case 1 - two features (mean and median); Case 2 - two features (mean and max. norm); Case 3 - three features (mean, median, and max. norm); Case 4 - four features (mean, median, max. norm, and std. dev.); Case 5 - five features (mean, median, max. norm, std. dev., and L1 norm); and Case 6 - eight features (mean, median, max. norm, std. dev., median absolute dev., mean absolute dev, L1 norm, and L2 norm).

To evaluate their influence, the classification accuracy for all faults implemented on Motor 2 under 80% motor loading and 45 Hz operating frequency using the stator current I2 for the six cases are tabulated in Table 5.8. It is found Case 6 has better accuracy for most cases, therefore, Case 6 is chosen as the features used in this paper.

5.6.3 Performance Evaluation of Trained Classifier Models

After the training, the performance of the trained classification models was evaluated in MATLAB through testing using a new set of testing data under two tests: test 1 with 80% training data and 20% testing data, and test 2 with 70% training data and 30% testing data. The training set contains labels of faults, but the testing contains new data without labels of faults. Both training and testing accuracy values for the two motors are provided in Table 5.9.

Table 5. 8: Influence of the number of Features on Classification accuracy for all Faults of Motor 2
(current I2 processed at 45Hz and 80% 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
Decision Trees	Fine Tree	47.8	85.6	85.6	98.9	98.9	98.3
	Medium Tree	46.1	85.6	85.6	98.9	98.9	98.3
	Coarse Tree	38.9	76.1	76.1	82.8	82.8	79.4
SVM	Linear SVM	33.9	82.8	75.6	91.1	91.1	91.1

	Quadratic SVM	45.6	72.2	81.1	98.9	93.9	97.2
	Cubic SVM	38.3	57.2	55.6	52.2	63.3	69.4
	Fine Gaussian SVM	43.9	76.7	77.8	93.3	90	90.6
	Medium Gaussian SVM	39.4	74.4	78.3	91.7	91.1	90.6
	Coarse Gaussian SVM	25	70	72.8	86.7	86.1	87.8
KNN	Fine KNN	40.6	60.6	50	55	62.8	70
	Medium KNN	38.9	50	41.1	39.4	44.4	44.4
	Cosine KNN	32.8	39.4	35.6	35.6	41.1	37.2
	Cubic KNN	37.8	48.3	41.1	38.3	43.9	42.8
	Weighted KNN	44.4	60	51.7	52.8	60.6	67.8
Ensemble	Boosted Trees	44.4	70.6	30.6	16.7	50	16.7
	Bagged Trees	46.7	85	82.8	88.9	91.1	100
	Subspace Discriminant	19.4	34.4	19.4	38.3	40.6	73.9
	Subspace KNN	28.9	52.8	85.6	100	100	100
	RUSBoosted Trees	45	73.3	31.1	16.7	50	16.7

Table 5. 9: Testing Performance of Trained Classifier Models with maximum accuracy for All Faults of Motor 1 and 2

Motor Name	Classification Method	Test 1 (80% training data, 20% testing data)		Test 2 (70% training data, 30% testing data)	
		Training Accuracy (%)	Testing Accuracy (%)	Training Accuracy (%)	Testing Accuracy (%)
Motor 1	Linear SVM	91.7	88.89	88.9	85.18
	Quadratic SVM	89.3	88.89	87.3	87.04
	Medium Gaussian SVM	89.6	86.11	84.9	83.33
	Subspace Discriminant	90	88.89	90	88.89
Motor 2	Fine Tree	98.3	97.2	98.2	96.3
	Medium Tree	98.3	97.2	98.2	96.3
	Linear SVM	90.5	88.89	90.1	88.89
	Quadratic SVM	97.2	94.44	94.4	94.44
	Fine Gaussian SVM	89.6	88.89	86.5	85.19
	Medium Gaussian SVM	91.7	91.67	90	88.89
	Bagged Trees	100	100	100	100
	Subspace KNN	100	100	100	100

5.7 Features Calculation Formulas Developed Through surface Fitting

In experiments, the motors were tested under six loadings (0%, 20%, 40%, 60%, 80%, and 100%) and six VFD output frequencies (45 Hz, 50 Hz, 55 Hz, 60 Hz, 65 Hz, and 70 Hz). In real life, the motor loading factor and the VFD output frequency can be values that are different from the testing data. Directly determination of features for those untested cases through DWT is not feasible.

To solve this problem, in this research, we propose to develop feature calculation formulas for untested cases through the surface fitting technique using the tested data. Surface fitting is a regressional process, where the relationship among a dependent variable and two independent variables is developed. In this paper, the motor operating frequencies and load factors are used as independent variables, the features are the function to be developed using the two independent variables.

5.7.1 Surface Fitting Method

To improve the effectiveness and accuracy of the developed equations, the least absolute residuals (LAR) robustness algorithm is used. It detects and cures outliers to follow the actual trend of the data set. In LAR, data points having absolute residual values higher than threshold are disregarded and thus, the main trend of the dataset is captured. LAR is an iterative method, and the equation used to estimate the least absolute deviation for LAR is [48]

$$\beta_{LAR} = arg\ min \sum_{i=1}^{n} |\varepsilon_i(\beta)| \qquad (4)$$

Where, β_{LAR} is the absolute deviation estimator, $\varepsilon_i(\beta)$ is the error, and n is the number of data samples. Table 5.10 shows surface fitting models of the DWT processed stator current I2 features, along with their R-square values for Motor 1 with a multi-fault (BF + 1BRB). In these models, x represents the operating frequency in Hz, y represents the percentage of loading (%), and f(x,y) represents the feature value. Polynomial 11 and polynomial 21 equations are chosen as the target functions. Six out of eight equations

have high R-square values, which indicate that the fitting equations follow the trend of actual measurement data. Only two equations for calculating features "mean" and "median" have low R-square values due to very small values, however, relative errors between the experimental based data and calculated data shown in Table 5.11 proves that the experimental based data match are also following the actual measurement. Fig. 5.13 shows the graphs of the eight features vs. the motor loading and operating frequency using the stator current I2 for Motor 1 with a multi-fault (BF + 1BRB). The dots are DWT processing results using experimental data; while the solid surface is determined by the surface fitting equations. Surface fitting equations for features of other types of faults can be determined using similar procedure.

Similarly, the surface fitting equations using z-axis vibration signal features are shown in Table 5.12 along with their R-square values for the same fault. Polynomial 11 and polynomial 21 models are adopted, where x represents the operating frequency in Hz, y represents the percentage of loading %), and f(x,y) represents the function to be developed to calculate new features. Relative errors between experimental based data and calculated data by curve fitting equations are shown in Table 5.13. Fig. 5.14 shows the graphs of the eight features vs. the motor operating frequency in Hz and loading factor in percentage. The dots are DWT processing results using experimental data, while the surface line is determined by the surface fitting equations.

Table 5. 10: Surface fitting models for Features using stator current I2 processed by DWT for Motor 1 with a multi-fault (BF + 1 BRB)

Features Name	Equation	R-square Values
Mean	$f(x,y) = 0.013 + 0.00014x + 9.6 * 10^{-6}y$	0.3217
Median	$f(x,y) = -0.010\,7 + 0.0012x + 0.0001y - 1.438 * 10^{-5}x^2 - 3.465 * 10^{-6}xy$	0.4112
Standard Deviation	$f(x,y) = 3.691 - 0.085x - 0.0047y + 0.00059x^2 + 0.0001168xy$	0.9933
Median Absolute Value	$f(x,y) = 3.261 - 0.074x - 0.0010y + 0.0005x^2 + 6.931 * 10^{-5}xy$	0.9964
Mean Absolute Value	$f(x,y) = 3.461 - 0.081x - 0.0044y + 0.0005x^2 + 0.0001xy$	0.9923

L1 Norm	$f(x,y) = 3.121E + 4 - 731.2x - 39.51y + 5.09x^2 + 0.9303xy$	0.9922
L2 Norm	$f(x,y) = 350.3 - 8.065x - 0.4626y + 0.055x^2 + 0.011xy$	0.9825
Maximum Norm	$f(x,y) = 6.816 - 0.1544x - 0.006y + 0.0011x^2 + 0.0001xy$	0.9714

Table 5. 11: Relative errors between experimental based data and calculated data for Motor 1 with a multi-fault (BF+1 BRB) processed by the stator current I2

Features Name	Experimental based data	Calculated Data	Error, %
Mean	0.006622	0.0065816	0.61
Median	0.007013	0.006983	0.43
Standard Deviation	0.9679	0.9575375	1.07
Median Absolute Value	1.037	1.03508	0.185
Mean Absolute Value	0.8483	0.8403249	0.94
L1 Norm	7635	7562.95	0.94
L2 Norm	91.82	90.868251	1.03
Maximum Norm	1.625	1.627175	-0.13

Table 5. 12: Regression models for Features using z-axis vibration signal processed by DWT for Motor 1, BF+1 BRB fault

Featured Name	Equation	R-square Value
Mean	$f(x,y) = 0.01926 - 6.133 * 10^{-5x} + 2.307 * 10^{-5y}$	0.8129
Median	$f(x,y) = -1.675 * 10^{-18} + 3.722 * 10^{-20x} + 4.181 * 10^{-25y}$	0.8053
Standard Deviation	$f(x,y) = 0.2703 - 0.007354x - 0.000776y + 6.608 * 10^{-5x^2} + 1.384 * 10^{-5xy}$	0.6308
Median Absolute Value	$f(x,y) = 0.04 - 2.782 * 10^{-19x} - 5.375 * 10^{-20y}$	0.8429
Mean Absolute Value	$f(x,y) = 0.177 - 0.005276x - 0.0002981y + 5.128 * 10^{-5x^2} + 5.439 * 10^{-6xy}$	0.8858
L1 Norm	$f(x,y) = 1900 - 56.64x - 4.408y + 0.5263x^2 + 0.07838xy$	0.8862
L2 Norm	$f(x,y) = 23.64 - 0.6239x - 0.07123y + 0.00565x^2 + 0.001293y$	0.6292
Maximum Norm	$f(x,y) = -3.9 + 0.2297x - 0.018y - 0.002x^2 + 0.0003201xy$	0.1257

Table 5. 13: Relative errors between experimental based data and calculated data (for Motor 1, BF+1 BRB fault, z-axis vibration signal)

Features Name	Experimental based data	Calculated data	Error, %
Mean	0.0182	0.0182	-0.021
Median	0	0	0
Standard Deviation	0.0629	0.06478	-3.07
Median Absolute Value	0.04	0.04	0
Mean Absolute Value	0.0397	0.04205	-5.857
L1 Norm	356.2	367.1475	-3.073
L2 Norm	6.21	6.405	-3.21
Maximum Norm	2	1.9958	0.21
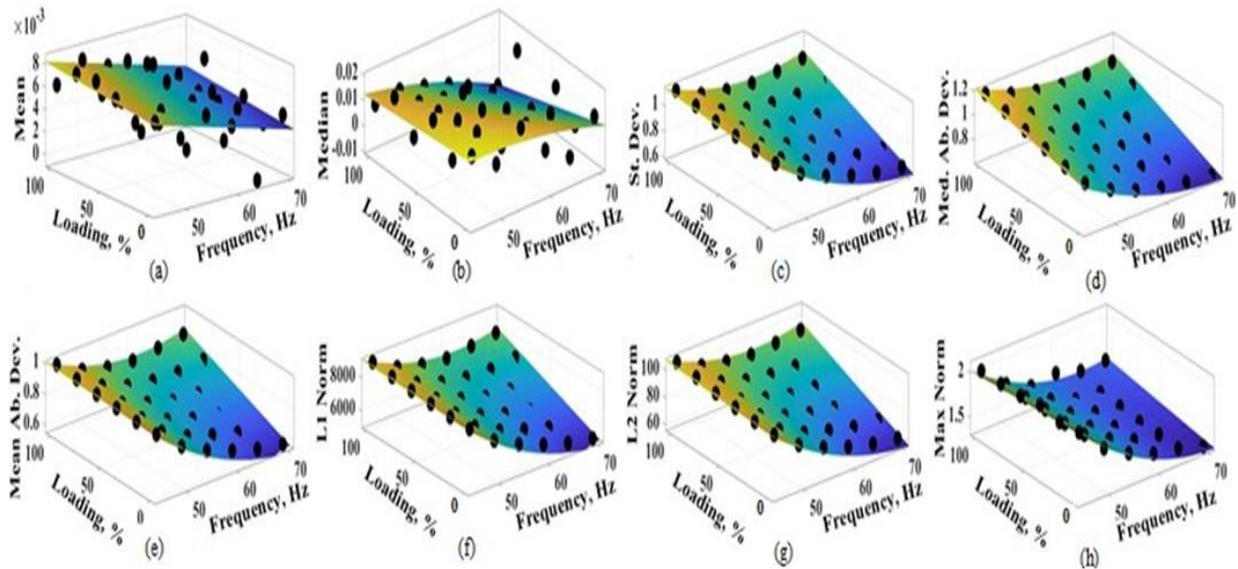
Fig. 5. 13. Surface fitting results for features for Motor 1 with a multi-fault (BF+1BRB) processed by the stator current I2: (a) mean, (b) median, (c) standard deviation, (d) median absolute value, (e) mean absolute value, (f) L1 norm, (g) L2 norm, and (h) maximum norm.
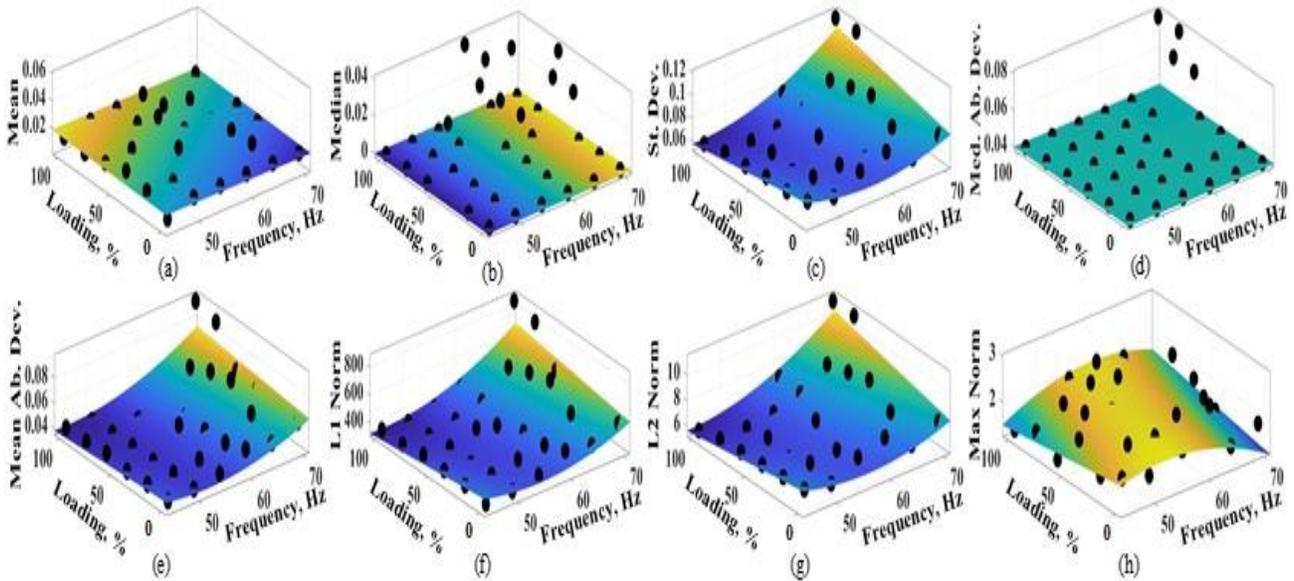
Fig. 5. 14. Surface fitting results for features for Motor 1 with a multi-fault (BF+1BRB) processed by the z-axis vibration signal: (a) mean, (b) median, (c) standard deviation, (d) median absolute value, (e) mean absolute value, (f) L1 norm, (g) L2 norm, and (h) maximum norm.

5.7.2 Machine Learning Results Using Fitting Equations

The features of Motor 1 for all healthy and faulty conditions processed by the stator current I_2 are calculated using the developed surface fitting equations for the following three cases, 90% at 64 Hz, 85% at 48 Hz, and 75% at 54 Hz, which have not been tested during experiments.

The results are shown in Fig. 15. Fig. 15 indicates similar performance to previous accuracy using tested cases, and thus, it proves that the surface fitting equations offer effective feature calculation of untested cases for induction motors fed by VFDs. The testing accuracy based on surface fitting methods for Motor 1 is also evaluated as shown in Table 5.14.
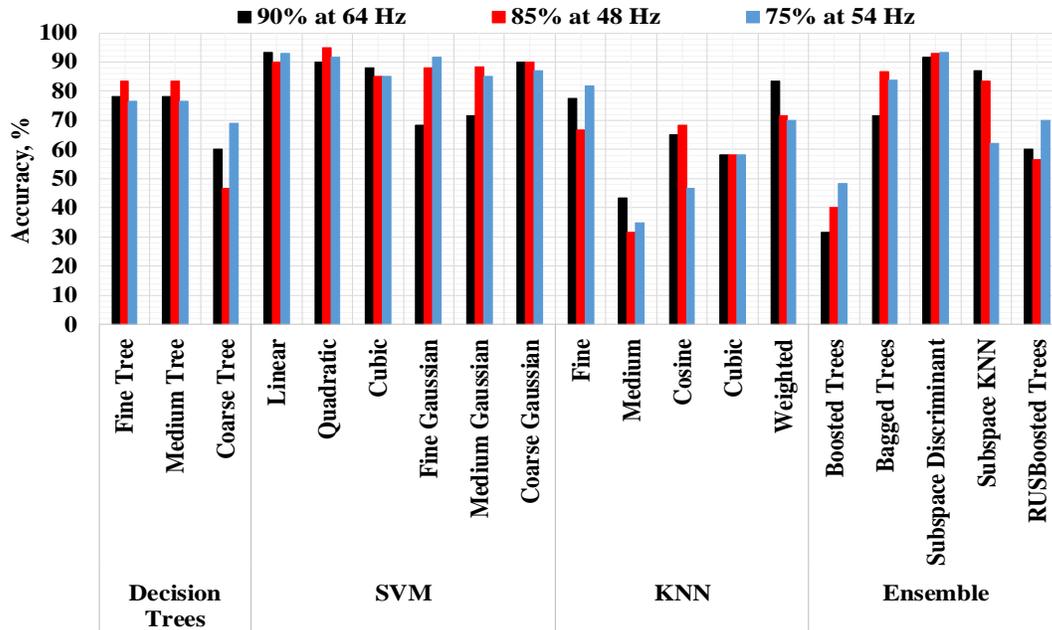
Fig. 5. 15. Classification accuracy for all faults of Motor 1 using features calculated by surface fitting equations for three untested cases (90% at 64 Hz, 85% at 48 Hz and 75% at 54 Hz) (processed using the stator current I2).

Table 5. 14: Testing Performance of Trained Classifier Models with maximum accuracy for Motor 1 after surface fitting processed data (using stator current I2)

Motor 1 Operating Conditions	Classification Method	Test 1 (80% training data, 20% testing data)		Test 2 (70% training data, 30% testing data)	
		Training Accuracy (%)	Testing Accuracy (%)	Training Accuracy (%)	Testing Accuracy (%)
85% at 48 Hz	Linear SVM	89.6	88.89	88.9	85.18
	Quadratic SVM	92.8	91.67	90	87.04
	Subspace Discriminant	91.7	88.89	90.5	88.89

5.8 Conclusion

In this paper, a machine learning based fault diagnosis method for induction motors fed by VFDs are proposed. Two identical induction motors are tested in the lab by using a VFD as the power supply. The tests were conducted considering different single- and multi-faults, VFD output frequencies, and motor loading factors. The experimental data in the form of one phase stator

current I2 and z-axis vibration signals are processed using DWT to extract features for machine learning. Eight features (mean, median, standard deviation, median absolute deviation, mean absolute deviation, L1 norm, L2 norm, and maximum norm) are extracted from the signal and verified to be the best feature combination.

Four families of classification algorithms in MATLAB Classification Learner toolbox with twenty classifiers are chosen to perform machine learning. It is found that among the twenty classifiers, 3 classifiers for Motor 1 and 8 classifiers for Motor 2 have accuracy above 90%. Among these high performance classifiers, two classifiers, Linear SVM and Medium Gaussian SVM, consistently appear for both motors, therefore, Linear SVM and Medium Gaussian SVM can serve as effective classifiers for fault diagnosis of induction motors fed by VFDs.

By comparing the classification accuracy for different types of faults under various operating conditions, it is found that the stator current performs better and offers higher accuracy than the vibration signal. Therefore, it is recommended that the motor stator current should be used for fault diagnosis of induction motors fed by VFDs.

The feature calculation formulas are developed through surface fitting using experimental data, these formulas are the function of the motor operating frequency in Hz and the motor loading factors in percentage. The purpose to develop these formulas is to calculate features for untested operating conditions in order to provide more comprehensive training to the chosen machine learning algorithm.

References:

[1] L. Sang Bin, G. B. Kliman, M. R. Shah, W. T. Mall, N. K. Nair, and R. M. Lusted, "An advanced technique for detecting inter-laminar stator core faults in large electric machines," *IEEE Trans. Ind. Appl.*, vol. 41, no. 5, pp. 1185–1193, Sep./Oct. 2005.

[2] H. Jafari and J. Poshtan, "Fault detection and isolation based on fuzzy-integral fusion approach," in *IET Sci., Meas. & Technol.*, vol. 13, no. 2, pp. 296-302, 3 2019.

[3] S. B. Jiang, P. K. Wong, R. Guan, Y. Liang and J. Li, "An Efficient Fault Diagnostic Method for Three-Phase Induction Motors Based on Incremental Broad Learning and Non-Negative Matrix Factorization," in *IEEE Access*, vol. 7, pp. 17780-17790, 2019.

[4] A. Khezzar, M. El Kamel Oumaamar, M. Hadjami, M. Boucherma, and H. Razik, "Induction motor diagnosis using line neutral voltage signatures," *IEEE Trans. Ind. Electron.*, vol. 56, no. 11, pp. 4581–4591, Nov. 2009.

[5] A. Ibrahim, M. El Badaoui, F. Guillet, and F. Bonnardot, "A new bearing fault detection method in induction machines based on instantaneous power factor," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, pp. 4252–4259, Dec. 2008.

[6] J. A. Corral-Hernandez and J. A. Antonino-Daviu, "Thorough validation of a rotor fault diagnosis methodology in laboratory and field soft-started induction motors," in *Chinese J. of Electr. Eng.*, vol. 4, no. 3, pp. 66-72, September 2018.

[7] R. N. Andriamalala, H. Razik, L. Baghli and F. Sargos, "Eccentricity Fault Diagnosis of a Dual-Stator Winding Induction Machine Drive Considering the Slotting Effects," in *IEEE Trans. on Ind. Electron.*, vol. 55, no. 12, pp. 4238-4251, Dec. 2008.

[8] F. Briz, M. W. Degner, J. M. Guerrero and P. Garcia, "Stator Windings Fault Diagnostics of Induction Machines Operated From Inverters and Soft-Starters Using High-Frequency Negative-Sequence Currents," in *IEEE Trans. on Ind. Appl.*, vol. 45, no. 5, pp. 1637-1646, Sept.-oct. 2009.

[9] F. Briz, M. W. Degner, P. Garcia and A. B. Diez, "High-Frequency Carrier-Signal Voltage Selection for Stator Winding Fault Diagnosis in Inverter-Fed AC Machines," in *IEEE Trans. on Ind. Electron.*, vol. 55, no. 12, pp. 4181-4190, Dec. 2008.

[10] B. Trajin, J. Regnier and J. Faucher, "Comparison Between Stator Current and Estimated Mechanical Speed for the Detection of Bearing Wear in Asynchronous Drives," in *IEEE Trans. on Ind. Electron.*, vol. 56, no. 11, pp. 4700-4709, Nov. 2009.

[11] R. J. Povinelli, J. F. Bangura, N. A. O. Demerdash and R. H. Brown, "Diagnostics of bar and end-ring connector breakage faults in polyphase induction motors through a novel dual track of time-series data mining and time-stepping coupled FE-state space modeling," in *IEEE Trans. on Energy Convers.*, vol. 17, no. 1, pp. 39-46, March 2002.

[12] S. M. A. Cruz and A. J. M. Cardoso, "Diagnosis of stator inter-turn short circuits in DTC induction motor drives," in *IEEE Trans. on Ind. Appl.*, vol. 40, no. 5, pp. 1349-1360, Sept.-Oct. 2004.

[13] G. Bossio, C. De Angelo, J. Solsona, G. O. García, and M. I. Valla, "Application of an additional excitation in inverter-fed induction motors for air-gap eccentricity diagnosis," *IEEE Trans. Energy Convers.*, vol. 21, no. 4, pp. 839–847, Dec. 2006.

[14] I. P. Georgakopoulos, E. D. Mitronikas, A. N. Safacas, and I. P. Tsoumas, "Detection of eccentricity in inverter-fed induction machines using wavelet analysis of the stator current," in *Proc. IEEE Power Electron. Spec. Conf.*, Rhodes Island, Greece, pp. 487–492, Jun. 15–19, 2008.

[15] B. Kim, K. Lee, J. Yang, S. B. Lee, E. J. Wiedenbrug and M. R. Shah, "Automated Detection of Rotor Faults for Inverter-Fed Induction Machines Under Standstill Conditions," in *IEEE Trans. on Ind. Appl.*, vol. 47, no. 1, pp. 55-64, Jan.-Feb. 2011.

[16] I. P. Georgakopoulos, E. D. Mitronikas and A. N. Safacas, "Detection of Induction Motor Faults in Inverter Drives Using Inverter Input Current Analysis," in *IEEE Trans. on Ind. Electron.*, vol. 58, no. 9, pp. 4365-4373, Sept. 2011.

[17] K. Teotrakool, M. J. Devaney and L. Eren, "Adjustable-Speed Drive Bearing-Fault Detection Via Wavelet Packet Decomposition," in *IEEE Trans. on Instrum. and Meas.*, vol. 58, no. 8, pp. 2747-2754, Aug. 2009.

[18] Y. Gritli, C. Rossi, D. Casadei, F. Filippetti and G. Capolino, "A Diagnostic Space Vector-Based Index for Rotor Elect. Fault Detection in Wound-Rotor Induction Machines Under Speed Transient," in *IEEE Trans. on Ind. Electron.*, vol. 64, no. 5, pp. 3892-3902, May 2017.

[19] A. Sapena-Baño *et al.*, "Condition monitoring of elect. machines using low computing power devices," *2014 Int. Conf. on Elect. Machines (ICEM)*, Berlin, pp. 1516-1522, 2014.

[20] A. Stefani, A. Bellini and F. Filippetti, "Diagnosis of Induction Machines' Rotor Faults in Time-Varying Conditions," in *IEEE Trans. on Ind. Electron.*, vol. 56, no. 11, pp. 4548-4556, Nov. 2009.

[21] J. Faiz and M. Ojaghi, "Instantaneous-Power Harmonics as Indexes for Mixed Eccentricity Fault in Mains-Fed and Open/Closed-Loop Drive-Connected Squirrel-Cage Induction Motors," in *IEEE Trans. on Ind. Electron.*, vol. 56, no. 11, pp. 4718-4726, Nov. 2009.

[22] K. M. Sousa, I. Brutkowski Vieira da Costa, E. S. Maciel, J. E. Rocha, C. Martelli and J. C. Cardozo da Silva, "Broken Bar Fault Detection in Induction Motor by Using Optical Fiber Strain Sensors," in *IEEE Sensors J.*, vol. 17, no. 12, pp. 3669-3676, 15 June15, 2017.

[23] M. F. Cabanas *et al*., "Unambiguous Detection of Broken Bars in Asynchronous Motors by Means of a Flux Measurement-Based Procedure," in *IEEE Trans. on Instrum. and Meas.*, vol. 60, no. 3, pp. 891-899, March 2011.

[24] J. Seshadrinath, B. Singh and B. K. Panigrahi, "Investigation of Vibration Signatures for Multiple Fault Diagnosis in Variable Frequency Drives Using Complex Wavelets," in *IEEE Trans. on Power Electron.*, vol. 29, no. 2, pp. 936-945, Feb. 2014.

[25] J. Seshadrinath, B. Singh and B. K. Panigrahi, "Broken rotor bar detection in variable frequency induction motor drives using wavelet energy based method," *2012 IEEE Int. Conf. on Power Electron., Drives and Energy Syst. (PEDES)*, Bengaluru , pp. 1-5, 2012.

[26] 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," in *IEEE Trans. on Ind. Appl.*, vol. 54, no. 3, pp. 2215-2224, May-June 2018.

[27] X. Huang, T. G. Habetler and R. G. Harley, "Detection of Rotor Eccentricity Faults in a Closed-Loop Drive-Connected Induction Motor Using an Artificial Neural Network," in *IEEE Trans. on Power Electron.*, vol. 22, no. 4, pp. 1552-1559, July 2007.

[28] M. Z. Ali, M. N. S. K. Shabbir, X. Liang, Y. Zhang, and T. Hu, "Machine Learning based Fault Diagnosis for Single- and Multi-Faults in Induction Motors Using Measured Stator Currents and Vibration Signals," *IEEE Transactions on Industry Applications*, pp. 1–1, 2019.

[29] O. Mohammed, N. Abed, and S. Ganu, "Modeling and Characterization of Induction Motor Internal Faults Using Finite-Element and Discrete Wavelet Transforms," *IEEE Transactions on Magnetics*, vol. 42, no. 10, pp. 3434–3436, 2006.

[30] N. Bessous, S. E. Zouzou, S. Sbaa, and W. Bentrah, "A comparative study between the MCSA, DWT and the vibration analysis methods to diagnose the dynamic eccentricity fault

in induction motors," *2017 6th International Conference on Systems and Control (ICSC)*, 2017.

[31] A. Mohanty and C. Kar, "Fault Detection in a Multistage Gearbox by Demodulation of Motor Current Waveform," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 4, pp. 1285–1297, 2006.

[32] 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.

[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] E. Hullermeier, "Experience-Based Decision Making: A Satisficing Decision Tree Approach," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 35, no. 5, pp. 641–653, 2005.

[35] X. Peng, H. Guo, and J. Pang, "Performance Analysis between Different Decision Trees for Uncertain Data," *2012 International Conference on Computer Science and Service System*, 2012.

[36] B. Scholkopf, K.-K. Sung, C. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik, "Comparing support vector machines with Gaussian kernels to radial basis function classifiers," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2758–2765, 1997.

[37] 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.

[38] 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.

[39] J. Seshadrinath, B. Singh, and B. K. Panigrahi, "Investigation of Vibration Signatures for Multiple Fault Diagnosis in Variable Frequency Drives Using Complex Wavelets," *IEEE Transactions on Power Electronics*, vol. 29, no. 2, pp. 936–945, 2014.

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

[41] 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.

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

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

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

[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] 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.

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

[48] M. N. S. K. Shabbir, X. Liang, W. Li, A. M. Le, and N. Khan, "A Data-Driven Voltage Control Approach for Grid-Connected Wind Power Plants," *2018 IEEE Industry Applications Society Annual Meeting (IAS)*, 2018.

Chapter 6

Conclusion

6.1 Summary

The main objective of this thesis is to develop fault diagnosis approaches for induction motors by employing machine learning and advanced signal processing techniques. Several diagnosis techniques are proposed in this thesis and verified to be effective for induction motors fed directly online or fed by VFDs. The main content of Chapters 3, 4 and 5 are summarized as follows:

In Chapter 3, a machine learning based fault diagnosis method for single- and multi-faults of induction motors fed directly online is proposed, developed, and validated using experimental data measured in the lab. Several conclusions are drawn in this chapter as follows: 1) The signal processing technique, MP or DWT can be addressed to extract features with comparable accuracy. 2) A quantitative comparison has been made and it is suggested that either stator current and vibration signals can be used to detect the same group of faults with a similar accuracy. 3) Five classifiers, Fine Gaussian SVM, fine KNN, weighted KNN, Bagged Trees, and subspace KNN are selected as suitable classifiers for induction motors fault diagnosis. 4) A novel curve fitting technique is developed to calculate features for the motors for which stator currents or vibration signals under certain loadings are not tested for a particular fault.

In Chapter 4, a robust fault diagnosis method is proposed for classifying various faults of induction motors based on the DWT processing results. In this chapter, stator currents of an 0.25 HP induction motor measured through an experimental test bench under healthy and faulty conditions and 100% loading are analyzed using the DWT for fault diagnosis. Two parameters are

evaluated, threshold and energy values, by the DWT processing. It is found that the threshold value for each decomposition level can serve as a good fault indicator of the motor.

In Chapter 5, a machine learning based fault diagnosis method considering different single- and multi-faults for induction motors fed by VFDs is proposed and verified. Four families of classification algorithms in MATLAB Classification Learner toolbox with twenty classifiers are chosen to perform machine learning. It is found that among the twenty classifiers, 3 classifiers for Motor 1 and 8 classifiers for Motor 2 have accuracy above 90%. Among these high performance classifiers, two classifiers, Linear SVM and Medium Gaussian SVM, consistently appear for both motors, therefore, Linear SVM and Medium Gaussian SVM can serve as effective classifiers for fault diagnosis of induction motors fed by VFDs. A quantitative comparison is also made and it is found that using the stator current performs better and offers higher accuracy than using the vibration signal for VFD-motor systems. Later, a novel surface fitting technique is developed to calculate features using experimental data for the motors for which stator currents or vibration signals under certain loadings and operating frequency are not tested for a particular fault, so the fault diagnosis can be conducted under any operating conditions.

6.2 Future Works

- 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.
- The experiment and result analysis are done considering the stator current and vibration signal. Other monitoring signals such as voltage, instant power, temperature, and torque may be also considered in the future.

List of Publications

(Since Fall 2017)

Refereed Journal Papers:

[1] **Mohammad Zawad Ali**, Md Nasmus Sakib Khan Shabbir, Xiaodong Liang, Yu Zhang, and Ting Hu, "Machine Learning based Fault Diagnosis for Single- and Multi-Faults in Induction Motors Using Measured Stator Currents and Vibration Signals," IEEE Transactions on Industry Applications (scheduled for May/June 2019 regular issue).

[2] Md Nasmus Sakib Khan Shabbir, **Mohammad Zawad Ali**, Xiaodong Liang and Muhammad Sifatul Alam Chowdhury, "A Probabilistic Approach Considering Contingency Parameters for Peak Load Demand Forecasting," IEEE Canadian Journal of Electrical and Computer Engineering, vol. 41, no. 4, pp. 224 - 233, Fall 2018.

[3] **Mohammad Zawad Ali**, and Xiaodong Liang, "Threshold Based Induction Motors Single- and Multi-Faults Diagnosis Using Discrete Wavelet Transform and Measured Stator Current Signal," submitted to IEEE Canadian Journal of Electrical and Computer Engineering (under review).

[4] **Mohammad Zawad Ali**, Md Nasmus Sakib Khan Shabbir, Shafi Md Kawsar Zaman, and Xiaodong Liang, "Machine Learning Based Fault Diagnosis for Single- and Multi-Faults for Induction Motors Fed by Variable Frequency Drives," submitted to IEEE Transactions on Industry Applications, pp. 1-14 (under review).

[5] Xiaodong Liang, **Mohammad Zawad Ali**, and Huaguang Zhang, "Fault Diagnosis for Induction Motors Using Finite Element Method – A Review," submitted to IEEE Transactions on Industry Applications, pp. 1-11 (under review).

Refereed Conference Papers:

[6] **Mohammad Zawad Ali**, Md Nasmus Sakib Khan Shabbir, Shafi Md Kawsar Zaman, and Xiaodong Liang, "Machine Learning Based Fault Diagnosis for Single- and Multi-Faults for Induction Motors Fed by Variable Frequency Drives," accepted by 54th IEEE Industry Applications Society (IAS) Annual Meeting, Baltimore, Maryland, United States, September 29th - October 3rd, 2019.

[7] Xiaodong Liang, **Mohammad Zawad Ali**, and Huaguang Zhang, "Fault Diagnosis for Induction Motors Using Finite Element Method – A Review," accepted by 54th IEEE Industry Applications Society (IAS) Annual Meeting, Baltimore, Maryland, United States, September 29th - October 3rd, 2019.

[8] **Mohammad Zawad Ali**, and Xiaodong Liang, "Induction Motor Fault Diagnosis Using Discrete Wavelet Transform," Proceedings of IEEE Canadian Conference of Electrical and Computer Engineering (CCECE) 2019, Edmonton, AB, Canada, May 5-8, 2019.

[9] **Mohammad Zawad Ali**, Md Nasmus Sakib Khan Shabbir, Xiaodong Liang, Yu Zhang, and Ting Hu, "Experimental Investigation of Machine Learning Based Fault Diagnosis for Induction Motors," Proceeding of IEEE Industry Applications Society (IAS) Annual Meeting, pp. 1-14, Portland, OR, USA, September 23 - 27, 2018.

[10] Yu Zhang, Ting Hu, Xiaodong Liang, **Mohammad Zawad Ali**, and Md Nasmus Sakib Khan Shabbir, "Fault Detection and Classification for Induction Motors using Genetic Programming," Proceeding of 22nd European Conference on Genetic Programming (EuroGP), Leipzig, Germany, 24-26 April, 2019.

[11] **Mohammad Zawad Ali**, Md Nasmus Sakib Khan Shabbir, Muhammad Sifatul Alam Chowdhury, Arko Ghosh, and Xiaodong Liang, "Regression Models of Critical Parameters Affecting Peak Load Demand Forecasting," Proceedings of the 31st Annual IEEE Canadian Conference on Electrical and Computer Engineering (CCECE 2018), pp. 1-4, Québec City, Québec, Canada, May 13-16, 2018.

[12] Md Nasmus Sakib Khan Shabbir, **Mohammad Zawad Ali**, Muhammad Sifatul Alam Chowdhury, and Xiaodong Liang, "A Probabilistic Approach for Peak Load Demand Forecasting," Proceedings of the 31st Annual IEEE Canadian Conference on Electrical and Computer Engineering (CCECE 2018), pp. 1-4, Québec City, Québec, Canada, May 13-16, 2018.

Non-Refereed Local IEEE Conference Papers (with Refereed Abstracts):

[13] **Mohammad Zawad Ali**, and Xiaodong Liang, "Characterization of Different Broken Rotor Bar Faults of Induction Motors by Orthogonal Matching Pursuit," Annual IEEE Newfoundland Electrical and Computer Engineering Conference (NECEC), pp. 1-5, St. John's, Canada, November 2018.

[14] **Mohammad Zawad Ali**, and Xiaodong Liang, "Optimization of High Efficiency AlxGa1-xAs & Ga0.47 In0.53AsyP1-y Bandgraded 2-J Tandem Solar Cell by Varying Mole Fraction of Tunnel Junction," Annual IEEE Newfoundland Electrical and Computer Engineering Conference (NECEC), pp. 1-5, St. John's, Canada, November 2017.