SIGNAL PROCESSING AND GRAPH-BASED SEMI-SUPERVISED LEARNING-BASED FAULT DIAGNOSIS FOR DIRECT ONLINE INDUCTION MOTORS

By

Shafi Md Kawsar Zaman

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 2020

St. John's Newfoundland and Labrador Canada

Abstract

In this thesis, fault diagnosis approaches for direct online induction motors are proposed using signal processing and graph-based semi-supervised learning (GSSL). These approaches are developed using experimental data obtained in the lab for two identical 0.25 HP three-phase squirrel-cage induction motors. Various electrical and mechanical single- and multi-faults are applied to each motor during experiments. Three-phase stator currents and three-dimensional vibration signals are recorded simultaneously in each experiment. In this thesis, Power Spectral Density (PSD)-based stator current amplitude spectrum analysis and one-dimensional Complex Continuous Wavelet Transform (CWT)-based stator current time-scale spectrum analysis are employed to detect broken rotor bar (BRB) faults. An effective single- and multi-fault diagnosis approach is developed using GSSL, where discrete wavelet transform (DWT) is applied to extract features from experimental stator current and vibration data. Three GSSL algorithms (Local and global consistency (LGC), Gaussian field and harmonic functions (GFHF), and greedy-gradient max-cut (GGMC)) are adopted and compared in this study. To enable machine learning for untested motor operating conditions, mathematical equations to calculate features for untested conditions are developed using curve fitting and features obtained from experimental data of tested conditions.

Acknowledgments

First and foremost, the author would like to convey sincere gratitude and respect to his supervisor Dr. Xiaodong Liang for her persistent support, scientific guidance, and constructive discussions throughout the development of this thesis.

The author profoundly appreciates the financial support from the Natural Science and Engineering Research Council of Canada (NSERC) Discovery Grant and the Graduate Fellowship from Memorial University. The author would also like to thank the Memorial University of Newfoundland for all kinds of support needed during this research.

Finally, the author wishes to thank his family members for their continual mental support and encouragement.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

IEEE	Institute of Electrical and Electronics Engineers
EPRI	Electric Power Research Institute
DWT	Discrete Wavelet Transform
SVM	Support Vector Machine
GSSL	Graph-Based Semi-Supervised Learning
SL	Supervised Learning
MCSA	Motor Current Signature Analysis
ANN	Artificial Neural Network
LGC	Local and Global Consistency
GFHF	Gaussian Field and Harmonic Functions
GGMC	Greedy-Gradient Max-Cut
LR	Label Ratio
UNB	Unbalanced shaft rotation
UV	Unbalanced Voltage
BF	Bearing Fault
BRB	Broken Rotor Bar
PSD	Power Spectral Density
CWT	Continuous Wavelet Transform
CMW	Complex Morlet Wavelet
FFT	Fast Fourier Transform
DFT	Discrete Fourier Transform
k-NN	k-Nearest Neighbors

List of symbols

y_i	The i^{th} Element of Dataset, i = 1, 2, 3,…, n
r	Residual
w_i	i^{th} element of the weights array
E	Set of Graph Edges
OvA	One-vs.-all
P	Number of Predictors
$\Psi(t)$	Mother Wavelet
f_b	Wavelet Bandwidth
f_c	Wavelet Center Frequency
X_l	Set of Labeled Inputs
X_u	Set of Unlabeled Inputs
$k(\cdot)$	Kernel Function
$Q(.)$	Objective Function
n	Total Number of samples

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

Chapter 1

Introduction

1.1 Background Information

Due to rugged construction, lower operating cost, and fewer maintenance requirements, induction motors are widely used in various industry sectors, such as centrifugal pumps, automotive, aerospace, conveyors, elevators, and packaging processes. Because of their robust build quality, they are often employed in hazardous locations, including oil fields, natural gas plants, and coal plants [1], [2].

Induction motors are subjected to various faults. The operational environment such as high ambient temperature, tasks performed, and installation issues may be combined to accelerate induction motor failure far sooner than the designed motor lifetime. If these faults are not diagnosed during impending stages, the motor can suffer severe damage, and production processes may come to a complete halt. Such unscheduled downtimes can increase the production cost to approximately twice the normal because of production shutdown, overhauling, and wastage of raw materials. Operators are under continual pressure to prevent this kind of unscheduled disruption and escalation in maintenance costs. Early detection of induction motor faults can attenuate this issue and ensure the reliable operation of critical industrial processes.

The faults of induction motors can be categorized, as shown in Fig. 1.1: rotor bar fault (5%), bearing fault (51%), stator winding fault (16%), shaft/coupling fault (2%), and the rest are classified as unknown faults or faults due to external reasons [3].

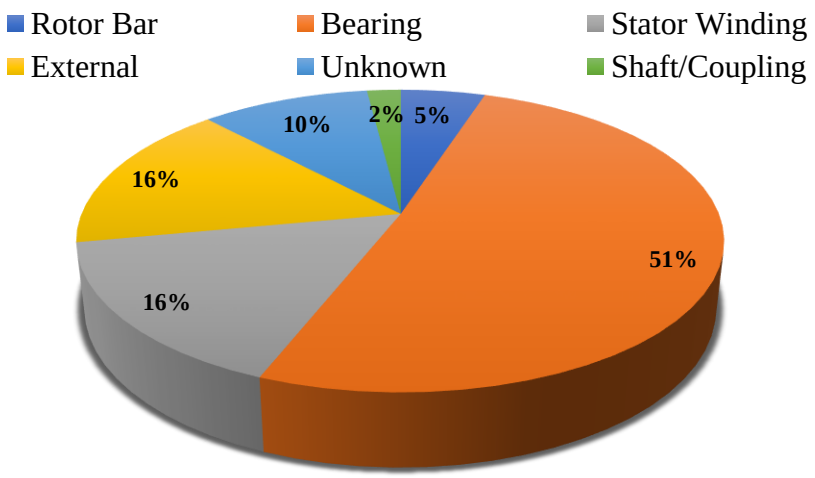


Fig. 1. 1. Different types of faults' distribution of an induction motor [3].

Similar statistical surveys are conducted by the IEEE Industrial Application Society (IAS) and electric power research institute (EPRI) on induction motor faults [4]. The percentages of different motor faults to the total number of faults are listed in Table 1.1.

Table 1. 1: Comparison of surveys from IEEE IAS and EPRI [4].

Major Components	IEEE IAS (%)	EPRI (%)
Rotor related	8	9
Bearing related	44	41
Stator related	26	36
Others	22	14

The most vulnerable parts of induction motors are bearing, stator winding, rotor bar, and shaft. Besides, faults also occur due to non-uniformity in the air gap between the stator-inner surface and the rotor-outer surface. Faults in induction motors can be categorized into electrical, mechanical, and environmental faults. Electrical faults include unbalanced supply voltage or current, under or over voltage or current, reverse phase sequence, earth fault, overload, and inter-turn short-circuit fault of the stator. Mechanical faults include broken rotor bar, mass unbalance, air gap eccentricity,

bearing damage, rotor winding failure, and stator winding failure. Environmental faults are mainly caused by the effect of ambient temperature and moisture. Vibrations due to installation defect and foundation defect also affect the performance of induction motors [2].

In induction motors, multiple faults may occur, and fault diagnosis becomes more challenging than single-fault [2], [5]. The effects of such faults in induction motors include unbalanced stator currents and voltages, oscillations in torque, reduction in efficiency and torque, overheating, and vibration. These faults can also increase the magnitude of specific harmonic components of currents and voltages [6].

Different failure modes call for comprehensive maintenance of induction motors. The maintenance can be categorized into: 1) scheduled maintenance, 2) breakdown maintenance, and 3) condition-based maintenance [7]. Scheduled maintenance includes planned check-up and repair of machines at a previously scheduled instance but results in extended downtime. This type of maintenance also requires an expert to precisely point out the defects, repairing or replacing the parts before the motor resumes working. In breakdown maintenance, the motor is allowed to run until it eventually wears out, which requires replacing that machine with an added cost to the process. Condition-based maintenance includes monitoring data, such as voltage and current spectrum, torque profile, during motor operation and taking necessary steps to prevent a fault at the developing stage to minimize the machine's downtime. Therefore, a condition-based monitoring system is desired because it requires less maintenance, lower cost, and drastically reduced production downtime [8].

In the age of industrial automation, significant progress has been made in computer-based data or signal acquisition and analysis, which has opened up a new pathway towards condition-based

maintenance as scheduled and breakdown maintenance are not sufficient as practicable choices anymore. The primary goal of the condition monitoring mechanism is to form a map between recorded signals and motor conditions to detect faults in incipient stages. By observing the aberrations in voltage, current, vibration, torque, and leakage flux, condition monitoring and fault diagnosis methods for induction motors can be developed.

1.2 Thesis Outline

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

Chapter 1

In Chapter 1, the importance of the research topic and background information are introduced; and objectives of the research are provided.

Chapter 2

In Chapter 2, a literature review is conducted for advanced signal processing and machine learning techniques for induction motor fault diagnosis.

Chapter 3

In Chapter 3, a stator current signature analysis method is proposed for fault diagnosis of squirrel-cage induction motor broken rotor bar (BRB) faults. Two different techniques are proposed and evaluated: Power Spectral Density (PSD) based stator current amplitude spectrum analysis; and one-dimensional Complex Continuous Wavelet Transform (CWT) based stator current time-scale spectrum analysis using Complex Morlet Wavelet (CMW). The performance of

the two techniques is compared using experimental stator current data measured in a lab for a 0.25 HP three-phase squirrel-cage induction motor. The stator current under healthy and faulty states of the motor was measured; the faults include one, two, and three BRBs. For 2 and 3 BRB faults, the holes were drilled on the rotor bars 90 degrees apart. The data of two motor loading conditions, 30%, and 85%, were used. It is found that CWT has better performance than PSD estimates for BRB fault detection. A version of this chapter has been published in proceedings of 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE).

Chapter 4

In Chapter 4, a GSSL-based fault diagnosis approach for direct online induction motors using both labeled and unlabeled data is proposed. Experimental data for two 0.25 HP induction motors under healthy and faulty conditions are used. Discrete Wavelet Transform (DWT) is employed to extract features from recorded stator current signals. Three GSSL algorithms (local and global consistency (LGC), Gaussian field and harmonic function (GFHF), and greedy-gradient max cut (GGMC)) are evaluated in this study, and GGMC shows superior performance over LGC and GFHF. They are also compared with a supervised learning algorithm, support vector machine (SVM). As induction motors often operate under variable loadings, curve fitting equations to calculate features for untested operating conditions are developed based on experimental data to enable machine learning for such untested conditions. A version of this chapter has been published in proceedings of 2020 Industrial and Commercial Power Systems Technical Conference.

Chapter 5

In Chapter 5, a GSSL-based fault diagnosis method for direct online induction motors is proposed using stator current and vibration signals. A 0.25 HP induction motor under healthy, single- and multi-fault states is tested in the lab. Three-phase stator currents and three-dimensional vibration signals of the motor are recorded simultaneously under steady-state operation in each test. Features for machine learning are extracted from the raw experimental stator current and vibration data using the DWT. Three GSSL algorithms, local and global consistency (LGC), Gaussian field and harmonic function (GFHF), and greedy-gradient max cut (GGMC), are used in the paper. It is found that both stator current and vibration signals perform well for one individual fault diagnosis using GSSL algorithms, but for classification of a combination of five different faults, the stator current outperforms the vibration signal significantly. Among the three GSSL algorithms, GGMC shows better classification results over LGC and GFHF for both stator current and vibration signals. A version of this chapter has been accepted by 2020 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE).

Chapter 6

In Chapter 6, research outcomes are summarized, and future work is recommended.

1.3 Objectives and Main Contributions of the Thesis

The objectives of this research focus on developing effective fault diagnosis methods for direct online induction motors using signal processing and graph-based semi-supervised learning (GSSL) techniques.

The main contributions of the thesis include:

1. In Chapter 3, the advanced signal processing-based broken rotor bar detection method for direct online induction motors is explored using experimental stator current signals. By comparing two signal processing methods, PSD and CWT, it is found that the CWT performs better in diagnosing BRB faults and distinguishing among different types of BRB faults.
2. In Chapter 4, an effective GGMC-based GSSL approach is proposed for induction motor single- and multi-fault diagnosis using experimental stator current signals. It is found the proposed method can effectively detect fault with high accuracy. Among the three GSSL algorithms, GGMC shows better classification results over LGC and GFHF.
3. In Chapter 5, a GSSL method is proposed for fault diagnosis of direct online induction motors using both stator current and vibration signals. It is found that both stator current and vibration signals perform well for the binary fault detection involving one individual fault; while the stator current outperforms vibration signals for the multiclass classification involving a combination of five different faults. Among the three GSSL algorithms, GGMC shows better classification results over LGC and GFHF for both stator current and vibration signals.

References

[1] X. Wen, “A hybrid intelligent technique for induction motor condition monitoring,” 2011.

[2] S. Karmakar, S. Chattopadhyay, M. Mitra, and S. Sengupta, *Induction motor fault diagnosis*, vol. 25. Springer, 2016.

[3] A. H. Bonnett and C. Yung, “Increased efficiency versus increased reliability,” *IEEE Industry Applications Magazine*, vol. 14, no. 1, pp. 29–36, 2008.

[4] P. Zhang, Y. Du, T. G. Habetler, and B. Lu, “A survey of condition monitoring and protection methods for medium-voltage induction motors,” *IEEE Transactions on Industry Applications*, vol. 47, no. 1, pp. 34–46, 2010.

[5] D. Basak, A. Tiwari, and S. Das, “Fault diagnosis and condition monitoring of electrical machines-A Review,” presented at the 2006 IEEE International Conference on Industrial Technology, 2006, pp. 3061–3066.

[6] S. Nandi, H. A. Toliyat, and X. Li, “Condition monitoring and fault diagnosis of electrical motors—A review,” *IEEE transactions on energy conversion*, vol. 20, no. 4, pp. 719–729, 2005.

[7] X. Liang and K. Edomwandekhoe, “Condition monitoring techniques for induction motors,” presented at the 2017 IEEE Industry Applications Society Annual Meeting, 2017, pp. 1–10.

[8] M. Ojaghi, M. Sabouri, and J. Faiz, “Diagnosis methods for stator winding faults in three-phase squirrel-cage induction motors,” *International transactions on electrical energy systems*, vol. 24, no. 6, pp. 891–912, 2014.

Chapter 2

Literature Review

Induction motors are the most commonly used electric machines in critical industrial applications. Although protective measures are employed, different electrical and mechanical faults still occur in induction motors, causing system breakdown and operational cost increase [1]. In the literature, three streams of work are reported in the literature for induction motors fault diagnosis, as shown in Fig. 2.1: 1) signature extraction-based approaches, 2) model-based approaches, and 3) knowledge-based approaches. The methods in bold in Fig. 2.1 are explored in this thesis.

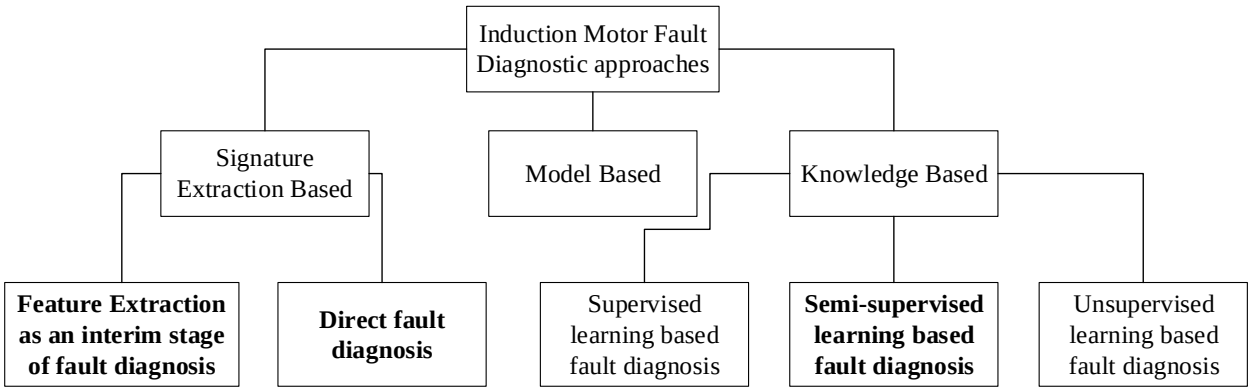


Fig. 2. 1. Induction motor fault diagnosis reported in the literature

2.1 Signature Extraction-Based Approaches

The signature extraction-based approaches are based on the analysis of fault signatures in time and frequency domains with respect to the recorded current, voltage, power, vibration, temperature, and acoustic signals [1]. The motor current signature analysis (MCSA) is one of the most popular techniques for online condition monitoring of induction motors [3]. MCSA is most successful in detecting broken rotor bars or end ring faults. Broken rotor bar faults are usually

diagnosed through characteristic harmonic sidebands [4], [5]. However, harmonic sidebands might also result from the design, construction, or load characteristics of induction motors. If these harmonic sidebands overlap with fault-related spectrum components, it is difficult to distinguish between a normal motor operating condition and a failure mode. Other advanced signal processing techniques including short-time Fourier transform (STFT), Wigner-Ville distribution (WVD), power spectral density (PSD), wavelet transform (WT) [6]–[8], multiple signal classification (MUSIC) [9], Hilbert transform [10], maximum covariance methods [11], and hybrid techniques (such as combining Wavelet and Hilbert transform [12], and homogeneity analysis with the Gaussian probability density function [13]) are reported in the literature for induction motor fault diagnosis. By choosing an appropriate signal processing approach, faults can be detected accurately. These signal processing techniques can also be utilized to extract features from recorded signals for machine learning [14].

2.2 Model-Based Approaches

The model-based fault diagnosis approaches require to model appropriate characteristics of induction motors under different fault modes [15]-[21]. Stator inter-turn short circuit fault diagnosis using a winding function approach is proposed in [16], but not suitable for industrial applications due to its inherent complexity. Fault diagnosis using negative sequence current [17], current vector radius [18], and multiple reference frame theory strategies [19] are proposed. Model-based approaches are the least explored in the literature as obtaining precise motor models for fault diagnosis is often infeasible [22].

2.3 Knowledge-Based Approaches

Knowledge-based approaches are data-driven machine learning-based techniques using data samples in online and offline applications [23].

The majority of the knowledge-based approaches uses supervised learning [24]. Artificial neural network (ANN) or other hybrid schemes of ANN is one of the most reported machine learning methods in the literature for fault diagnosis. The ANN methods include neural network with analytical redundancy method [25], Park's vector pattern learning [26], convolutional discriminative feature learning [27], fuzzy logic [28]. However, neural networks for online fault diagnosis of induction motors are mostly unsupervised [29], [30]. Other supervised machine learning approaches for induction motor fault diagnosis, such as support vector machine (SVM), k- nearest neighbor (KNN), ensemble and decision tree, are reported in the literature [24].

Obtaining the dataset is not a difficult task as condition data are continuously monitored in industries, but the process of labeling the collected data samples requires expert intervention. Updating the trained machine learning model for new data is also computationally expensive [31].

To alleviate this issue, semi-supervised learning methods are adopted that use only a few numbers of labeled samples along with a large number of unlabeled samples to construct a classification model [32]. These few labeled data can extract useful information from unlabeled data without any human involvement for labeling. Several diagnostic systems are designed based on semi-supervised learning strategies to improve efficiency and accuracy of the fault diagnostic systems.

Semi-supervised smooth alpha layering, semi-supervised label consistent dictionary learning, semi-supervised deep learning, semi-supervised multiple association layers network, and semi-supervised learning with manifold regularization based machine fault diagnostic methods are proposed in [23], [33]–[36].

Graph-based semi-supervised learning is a promising new area in the semi-supervised learning domain for active propagation of a limited number of initial labels to a large amount of unlabeled data. Fault detection and classification in PV arrays using GSSL is proposed in [31]. In [37], three different semi-supervised learning algorithms, local and global consistency (LGC), the Gaussian random field (GRF) method, and the graph transduction via alternating minimization (GTAM), are used and compared based on simulated and real benchmark datasets. Residential non-intrusive load monitoring using multi-label GSSL is proposed in [38].

This thesis focuses on two prime components for induction motor fault diagnosis: 1) an approach solely based on comparison of two distinctive signal processing methods, and 2) a proposed novel approach using graph-based semi-supervised learning.

References

[1] X. Liang and K. Edomwandekhoe, “Condition monitoring techniques for induction motors,” presented at the 2017 IEEE Industry Applications Society Annual Meeting, 2017, pp. 1–10.

[2] X. Dai and Z. Gao, “From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis,” *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2226–2238, 2013.

[3] J.-H. Jung, J.-J. Lee, and B.-H. Kwon, “Online diagnosis of induction motors using MCSA,” *IEEE Transactions on Industrial Electronics*, vol. 53, no. 6, pp. 1842–1852, 2006.

[4] M. E. H. Benbouzid and G. B. Kliman, “What stator current processing-based technique to use for induction motor rotor faults diagnosis?,” *IEEE Transactions on Energy Conversion*, vol. 18, no. 2, pp. 238–244, 2003.

[5] F. Duan and R. Zivanovic, “Induction motor fault diagnostics using global optimization algorithm,” presented at the 2009 Australasian Universities Power Engineering Conference, 2009, pp. 1–5.

[6] S. M. K. Zaman, H. U. M. Marma, and X. Liang, “Broken Rotor Bar Fault Diagnosis for Induction Motors Using Power Spectral Density and Complex Continuous Wavelet Transform Methods,” presented at the 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), 2019, pp. 1–4.

[7] J. CusidÓCusido, L. Romeral, J. A. Ortega, J. A. Rosero, and A. G. Espinosa, “Fault detection in induction machines using power spectral density in wavelet decomposition,” *IEEE Transactions on Industrial Electronics*, vol. 55, no. 2, pp. 633–643, 2008.

[8] A. Saghafinia, S. Kahourzade, A. Mahmoudi, W. P. Hew, and M. N. Uddin, “On line trained fuzzy logic and adaptive continuous wavelet transform based high precision fault detection of IM with broken rotor bars,” presented at the 2012 IEEE Industry Applications Society Annual Meeting, 2012, pp. 1–8.

[9] 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.

[10] 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.

[11] A. Bellini, G. Franceschini, and C. Tassoni, “Monitoring of induction machines by maximum covariance method for frequency tracking,” *IEEE Transactions on Industry Applications*, vol. 42, no. 1, pp. 69–78, 2006.

[12] 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.

[13] 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 Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1760–1769, 2017.

[14] F. Duan and R. Živanović, “Condition monitoring of an induction motor stator windings via global optimization based on the hyperbolic cross points,” *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1826–1834, 2014.

[15] A. Bellini, F. Filippetti, C. Tassoni, and G.-A. Capolino, “Advances in diagnostic techniques for induction machines,” *IEEE Transactions on industrial electronics*, vol. 55, no. 12, pp. 4109–4126, 2008.

[16] G. M. Joksimovic and J. Penman, “The detection of inter-turn short circuits in the stator windings of operating motors,” *IEEE Transactions on Industrial electronics*, vol. 47, no. 5, pp. 1078–1084, 2000.

[17] R. M. Tallam, T. G. Habetler, and R. G. Harley, “Transient model for induction machines with stator winding turn faults,” *IEEE Transactions on Industry Applications*, vol. 38, no. 3, pp. 632–637, 2002.

[18] S. M. Cruz and A. M. Cardoso, “Stator winding fault diagnosis in three-phase synchronous and asynchronous motors, by the extended Park’s vector approach,” *IEEE Transactions on industry applications*, vol. 37, no. 5, pp. 1227–1233, 2001.

[19] S. M. Cruz and A. M. Cardoso, “Multiple reference frames theory: A new method for the diagnosis of stator faults in three-phase induction motors,” *IEEE Transactions on Energy Conversion*, vol. 20, no. 3, pp. 611–619, 2005.

[20] X. Chang, V. Cocquempot, and C. Christophe, “A model of asynchronous machines for stator fault detection and isolation,” *IEEE transactions on industrial electronics*, vol. 50, no. 3, pp. 578–584, 2003.

[21] S. Bachir, S. Tnani, J.-C. Trigeassou, and G. Champenois, “Diagnosis by parameter estimation of stator and rotor faults occurring in induction machines,” *IEEE Transactions on Industrial Electronics*, vol. 53, no. 3, pp. 963–973, 2006.

[22] K. Kim and A. Parlos, “Model-based fault diagnosis of induction motors using non-stationary signal segmentation,” *Mechanical Systems and Signal Processing*, vol. 16, no. 2–3, pp. 223–253, 2002.

[23] R. Razavi-Far, E. Hallaji, M. Farajzadeh-Zanjani, and M. Saif, “A semi-supervised diagnostic framework based on the surface estimation of faulty distributions,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1277–1286, 2018.

[24] M. Z. Ali, M. N. S. K. Shabbir, S. M. K. Zaman, and X. Liang, “Single-and Multi-Fault Diagnosis Using Machine Learning for Variable Frequency Drive-Fed Induction Motors,” *IEEE Transactions on Industry Applications*, vol. 56, no. 3, pp. 2324–2337, 2020.

[25] H. Su and K. T. Chong, “Induction machine condition monitoring using neural network modeling,” *IEEE Transactions on Industrial Electronics*, vol. 54, no. 1, pp. 241–249, 2007.

[26] 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 Transactions on industry applications*, vol. 36, no. 3, pp. 730–735, 2000.

[27] W. Sun, R. Zhao, R. Yan, S. Shao, and X. Chen, “Convolutional discriminative feature learning for induction motor fault diagnosis,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1350–1359, 2017.

[28] S. Altug, M.-Y. Chen, and H. J. Trussell, “Fuzzy inference systems implemented on neural architectures for motor fault detection and diagnosis,” *IEEE transactions on industrial electronics*, vol. 46, no. 6, pp. 1069–1079, 1999.

[29] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, “Real-time motor fault detection by 1-D convolutional neural networks,” *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067–7075, 2016.

[30] M. Seera and C. P. Lim, “Online motor fault detection and diagnosis using a hybrid FMM-CART model,” *IEEE transactions on neural networks and learning systems*, vol. 25, no. 4, pp. 806–812, 2013.

[31] Y. Zhao, R. Ball, J. Mosesian, J.-F. de Palma, and B. Lehman, “Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays,” *IEEE Transactions on Power Electronics*, vol. 30, no. 5, pp. 2848–2858, 2014.

[32] D. Wu, X. Luo, G. Wang, M. Shang, Y. Yuan, and H. Yan, “A highly accurate framework for self-labeled semisupervised classification in industrial applications,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 909–920, 2017.

[33] W. Jiang, Z. Zhang, F. Li, L. Zhang, M. Zhao, and X. Jin, “Joint label consistent dictionary learning and adaptive label prediction for semisupervised machine fault classification,” *IEEE Transactions on Industrial Informatics*, vol. 12, no. 1, pp. 248–256, 2015.

[34] R. Razavi-Far *et al.*, “Information fusion and semi-supervised deep learning scheme for diagnosing gear faults in induction machine systems,” *IEEE Transactions on Industrial Electronics*, vol. 66, no. 8, pp. 6331–6342, 2018.

[35] K. Zhang, B. Tang, Y. Qin, and L. Deng, “Fault diagnosis of planetary gearbox using a novel semi-supervised method of multiple association layers networks,” *Mechanical Systems and Signal Processing*, vol. 131, pp. 243–260, 2019.

[36] J. Yuan and X. Liu, “Semi-supervised learning and condition fusion for fault diagnosis,” *Mechanical Systems and Signal Processing*, vol. 38, no. 2, pp. 615–627, 2013.

[37] T. Jebara, J. Wang, and S.-F. Chang, “Graph construction and b-matching for semi-supervised learning,” presented at the Proceedings of the 26th annual international conference on machine learning, 2009, pp. 441–448.

[38] D. Li and S. Dick, “Residential household non-intrusive load monitoring via graph-based multi-label semi-supervised learning,” *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4615–4627, 2018.

Chapter 3

Broken Rotor Bar Fault Diagnosis for Induction Motors Using Power Spectral Density and Complex Continuous Wavelet Transform Methods

Shafi Md Kawsar Zaman, *Student Member*, IEEE, Hla U May Marma, *Student Member*, IEEE, 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 in Proceedings of 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE). Shafi Md Kawsar Zaman co-authored this paper under the supervision of Dr. Xiaodong Liang. Shafi's contributions are listed as follows:

- Performed literature searches required for signal processing-based induction motor fault diagnosis.
- Conducted signal processing and evaluated the two proposed methods using experimented data.
- Examined the results.
- Involved in the writing of the paper draft as the first author.

Dr. Xiaodong Liang set up experimental plans, provided continuous technical guidance, checked the results, and modified the manuscript. Hla U May Marma helped with the literature review and involved with the writing of "Introduction" 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- Induction motors are widely used in various industrial sectors, fault diagnosis of induction motors are critical to prevent equipment failure and production downtime. In this paper, a stator current signature analysis method is proposed for squirrel cage induction motors' broken rotor bar (BRB) fault diagnosis. Two different techniques are implemented: Power Spectral Density (PSD) based stator currents' amplitude spectrum analysis; and one dimensional Complex Continuous Wavelet Transform (CWT) based stator currents' time-scale spectrum analysis using Complex Morlet Wavelet (CMW). The performance of the two techniques are compared using experimental stator current data measured in a lab for a 0.25 HP induction motor. The stator current under healthy and faulty states of the motor were measured, the faults include one, two and three BRBs. For 2 and 3 BRB faults, the holes were drilled on the rotor bars 90 degree apart. Two loading conditions of the motor were used during the measurement, 30% and 85%. It is found that the CWT has better performance than the PSD estimates for the BRB fault detection.

Keywords- Fault diagnosis, Complex continuous wavelet transform, Complex Morlet wavelet, Power spectral density estimate.

3.1 Introduction

Induction motors are most widely used electrical machines. Fault diagnosis is critical especially for machines operating in important industry processes or harsh environment. The squirrel cage induction motor consists of rotor bars and a shorted end ring. When a rotor bar is damaged or partially cracked, it is known as a broken rotor bar fault. The main causes of broken rotor bars of

an induction motor include manufacturing defects, fatigue of the metal due to thermal stresses, mechanical stress resulting from bearing faults, frequent starts and stops of the motor [1]. For induction motors, the broken rotor bar (BRB) faults are about 10% among various faults, such as bearing fault, stator winding inter-turn fault and insulation failure, and air-gap eccentric fault [2][3]. Therefore, the effective BRB fault diagnosis is important for the operation of induction motors. Advanced signal processing methods, such as Hilbert transform, Wigner-Ville distribution, stochastic resonance, and wavelet transform, have been applied in fault diagnosis of induction motors [4]-[8].

The most popular technique for BRB fault detection of induction motors is the motor current signature analysis (MCSA). The power spectrum estimation, fast Fourier transform (FFT), envelope spectrum analysis have been used to detect BRB faults. However, the traditional FFT cannot process non-stationary signals [5]. To overcome such issues, the Power Spectral Density (PSD) analysis, which analyzes the spectrogram based on the PSD estimates, is applied for the motor fault detection in [4]. Considering non-stationary and non-linear nature of the motor fault current signals, Wavelet analyzer can serve as a time-frequency analyzing tool, as it allows multi-resolution analysis in both time and frequency domains. There are basically two types of wavelet transforms: Discrete Wavelet Transform (DWT) and Continuous Wavelet Transform (CWT) [6]-[8]. The CWT algorithm is used to analyze different BRB faults in [6]. A new signal processing technique, known as complex CWT, has been used in [7]-[8], as the complex CWT has no endpoint effect and has small computation requirements, so it can enhance the accuracy of fault diagnosis. The bearing fault detection of induction motors is investigated by complex CWT in [8], but the BRB fault is not analyzed.

In this paper, two techniques are implemented for fault diagnosis of BRB faults in induction motors using experimental stator current signals: 1) power spectral density (PSD) analysis; and 2) one dimensional Complex CWT using Complex Morlet Wavelet (CMW). The experiments were conducted in a lab for a 0.25 HP squirrel cage induction motor. The healthy and BRB faults were tested for the motor. One, two and three BRB faults were created by drilling a hole on the motor bar. The size of the hole is the diameter of 4.2 mm and the depth of 18 mm. For two or three BRB faults, the bars with a hole on them were 90 mechanical degrees apart. The stator currents at three phases were measured using a power quality analyzer. The measured stator current at phase A is used as signal in this paper.

The paper is organized as follows: in Section 3.2, the principle of the CMW based CWT and the PSD algorithms are described; In Section 3.3, the measured stator current signals are analyzed using the two algorithms; Conclusions are drawn in Section 3.4.

3.2 Principle of Complex CWT and PSD

3.2.1 Complex Morlet Wavelet Based CWT Algorithm

The CWT algorithm is used to decompose a signal into wavelets. To compute each part of the time-domain signal individually, a window function, known as the mother wavelet, is used in the CWT. The mother wavelet $\Psi(t)$ satisfies (1).

$$\int_{-\infty}^{\infty} \Psi(t)dt = 0 \tag{1}$$

The CWT or the coefficient of the wavelet can be expressed by

$$CWT(b,a) = \int f(t){\Psi^*}_{b,a}(t)dt \tag{2}$$

The function f(t) is decomposed into a set of basic functions $\Psi^*\left(\frac{t-b}{a}\right)$, known as wavelets, which are generated from the mother wavelet by scaling and translation.

$$\Psi_{b,a}(t) = |a|^{-\frac{1}{2}}\ \Psi\left(\frac{t-b}{a}\right) \tag{3}$$

Where, a is the scale factor or window length, and b is the translation factor. The factor $|a|^{-\frac{1}{2}}$ is for energy normalization across different scales, and ${\Psi^*}_{b,a}(t)$ is conjugate to the mother wavelet function $\Psi_{b,a}(t)$. If the wavelet function in (2) is complex, it is defined as a complex wavelet transformation. The Complex Morlet Wavelet (CMW) is particularly useful for fault diagnosis due to its smallest time-frequency window area, smoothness and harmonic-like waveform [7][9]. The CMW is able to separate amplitude and phase information of the signal. The complex Morlet mother wavelet is given as follows [7]:

$$\Psi_0(t) = \frac{1}{\sqrt{\pi f_b}} exp\ (2i\pi f_c t)\ \exp(-\frac{t^2}{f_b}) \tag{4}$$

Where, f_b is the wavelet bandwidth and f_c is the wavelet center frequency.

3.2.2 Power Spectral Density Algorithm

Fourier Transform of the auto-correlation function is defined as the PSD of a discrete time process [10]. The periodogram is a PSD estimator of a complex discrete-time wide sense stationary random process $x\ [n]$, which is defined as follow:

$$P_x(f) = N^{-1}|\sum_{n=0}^{N-1} x[n]\exp(\frac{-j2\pi fn}{F_s})\ |^2 \tag{5}$$

The frequency resolution is equal to the inverse of the signal acquisition duration. The periodogram is usually implemented using the FFT algorithm since it can rapidly compute the discrete Fourier transform (DFT) [11].

When a BRB fault occurs in induction motors, harmonic sidebands around the supply frequency can be expected in the phase current power spectrum. The sideband situated on the left of the supply frequency is due to electrical or magnetic rotor asymmetry caused by BRBs, while the sideband situated on the right is due to speed ripples [12]. The amplitudes of the sidebands are affected by the position of BRBs, speed and loading conditions. The sidebands might be observed when the motor has no BRB faults but with rotor ellipticity or shaft misalignment, which could induce rotor asymmetry. However, the sideband amplitudes in these cases are typically smaller than that produced by BRB faults [13].

3.3 Analysis Using Complex CWT and PSD

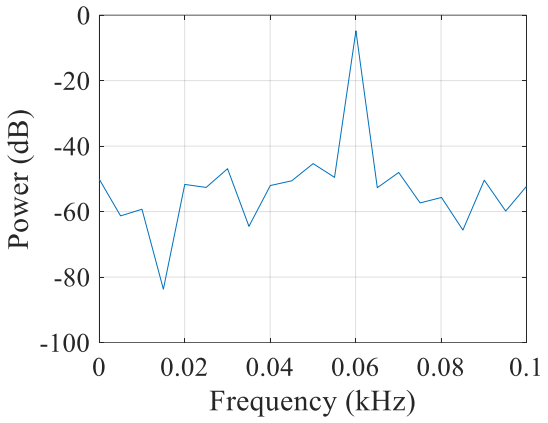
3.3.1 Periodogram PSD Estimates

Fourier coefficients [14] based PSD can be used for BRB fault detection in induction motors. The motor stator current signature is analyzed during normal operation. Since the sideband situated on the left of the supply frequency is caused by electrical or magnetic rotor asymmetry under a BRB fault condition, so the amplitudes of the left sidebands are investigated. The amplitudes in decibels (dB) at the frequencies of 30 Hz and 60 Hz are considered in the analysis, because 30 Hz (the first order sideband) is the most significant frequency component in the periodogram, which is situated on the left of the supply frequency of 60 Hz.

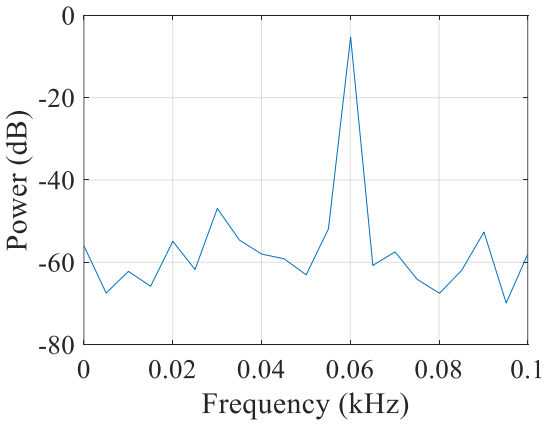
To perform PSD analysis, 3072 sample points of the measured stator current signal are taken for the healthy and faulty cases with 1 BRB, 2 BRB and 3 BRB faults. The sampling frequency of 15.38 kHz is used for the PSD analysis in all cases. Figs. 3.1 and 3.2 show the periodogram PSD estimates of the healthy and faulty motors at two different motor loading factors, 30% and 85%. The power amplitudes in dB of a healthy motor at 60 Hz are −4.80 dB and −2.38 dB for 30% and

85% loading, respectively. At the left sideband frequency of 30 Hz, the power amplitudes in dB are −46.91 dB and −54.76 dB for 30% and 85% loading, respectively. The amplitudes for the motor with 1 BRB, 2 BRB and 3 BRB faults with respect to 60 Hz and 30 Hz are tabulated in Table 3.1.

In Table 3.1, it can be seen that the differences in amplitudes between 60 Hz and 30 Hz frequency components for the healthy motor, 1 BRB, 2 BRB, and 3 BRB faulty motor under 30% loading are −42.11 dB, −41.64 dB, −41.18 dB and −35.95 dB, respectively. Similarly, the difference in amplitudes between 60 Hz and 30 Hz frequency components for healthy, 1 BRB, 2 BRB, and 3 BRB conditions under 85% loading are −52.38 dB, −47.99 dB, −43.80 dB, and −38.55 dB, respectively.



(a)



(b)

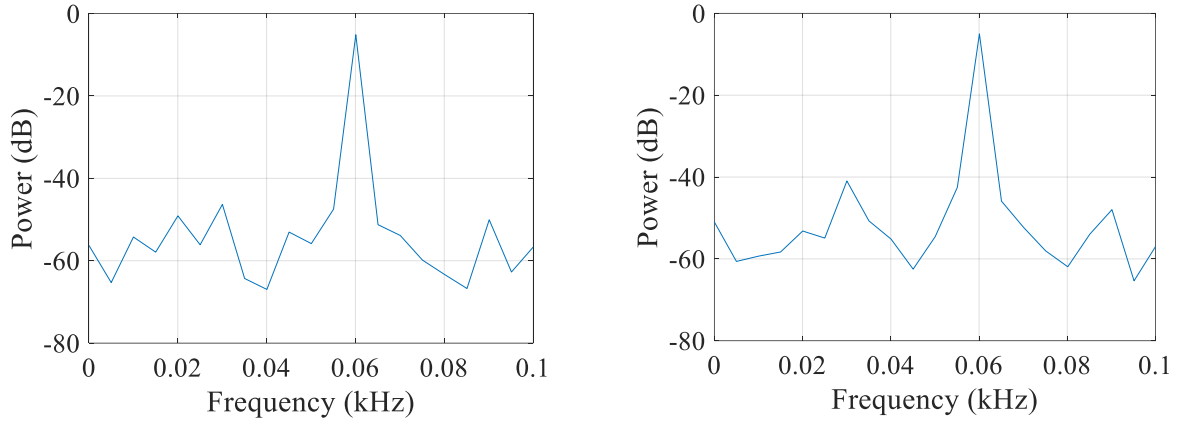


Fig. 3. 1. PSD analysis of the current signatures at 30% motor loading: (a) Healthy; (b) 1 BRB; (c) 2 BRB; (d) 3 BRB.

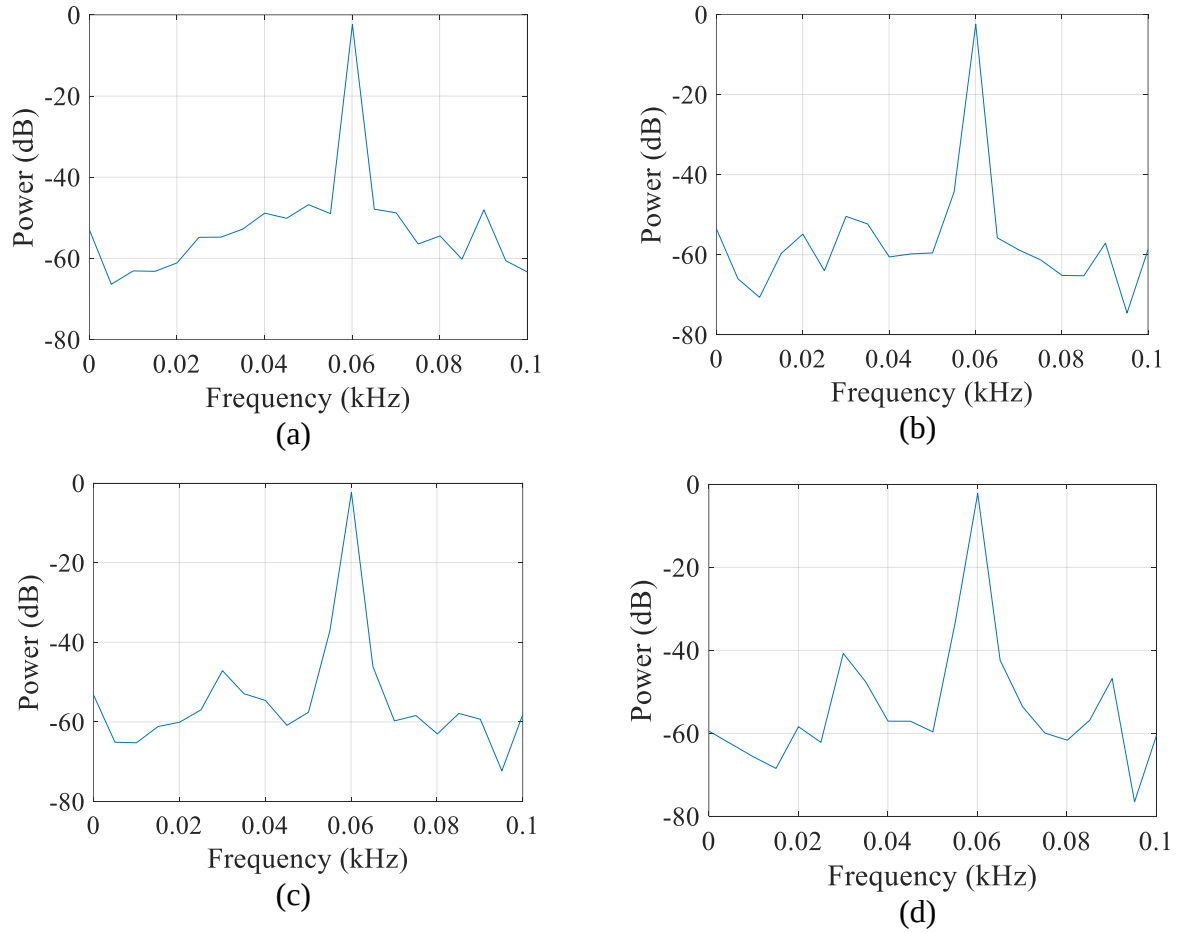


Fig. 3. 2. PSD analysis of the current signatures at 85% motor loading: (a) Healthy; (b) 1 BRB; (c) 2 BRB; (d) 3 BRB.

Table 3. 1: Power Amplitudes Extracted from PSD Estimates

Motor loading, %	Amplitude (dB) (Healthy motor at 60 Hz)	Left sideband amplitude (dB) (Healthy motor at 30 Hz)	Amplitude (dB) (Motor with 1 BRB at 60 Hz)	Left sideband amplitude (dB) (Motor with 1 BRB at 30 Hz)	Amplitude (dB) (Motor with 2 BRB at 60 Hz)	Left sideband amplitude (dB) (Motor with 2 BRB at 30 Hz)	Amplitude (dB) (Motor with 3 BRB at 60 Hz)	Left sideband amplitude (dB) (Motor with 3 BRB at 30 Hz)
30	–4.80	–46.91	–5.29	–46.93	–5.15	–46.33	–5.02	–40.97
85	–2.38	–54.76	–2.43	–50.42	–2.34	–47.14	–2.19	–40.74

It is found that for a light loading condition (30%), the differences in amplitudes between the 60 Hz and 30 Hz components are very close to one another for healthy and faulty cases. For a robust fault diagnosis, the differences are desired to increase with the increase of the fault severity. For a 30% loading, although the amplitude differences allow successful identification between healthy and faulty conditions, distinguishing the severity of faults can be difficult. For a heavier loading (85%), there is a significant increase in the difference in amplitudes between 60 Hz and 30 Hz components for healthy, 1 BRB, 2 BRB and 3 BRB conditions, so the PSD analysis can detect BRB faults and differentiate them according to the severity of the fault.

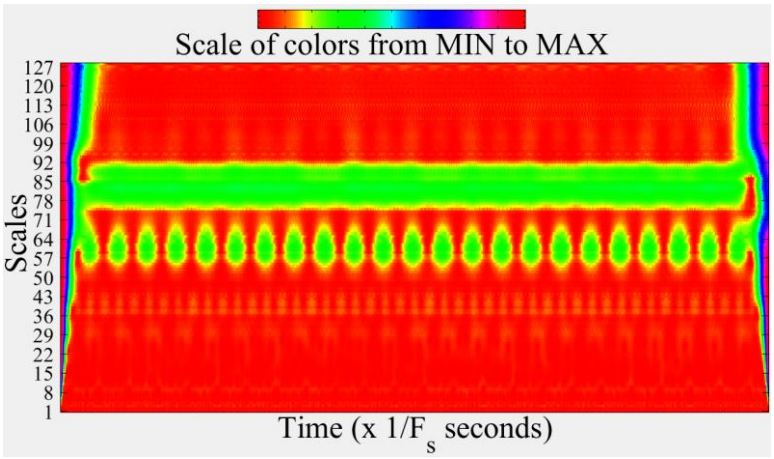
3.3.2 One dimensional (1-D) Complex CWT

One dimensional complex CWT is carried out using the Wavelet Analyzer Toolbox of MATLAB 2018b to visualize the time-scale scalograms of the measured stator current signals. The current signals are decomposed into 128 scales with a step size of 1, using the CMW as the mother wavelet. The bandwidth parameter is selected as ‘1-1.5’ for a better resolution of the time-scale representation of the complex CWT magnitude coefficients. The time-scale plots of the

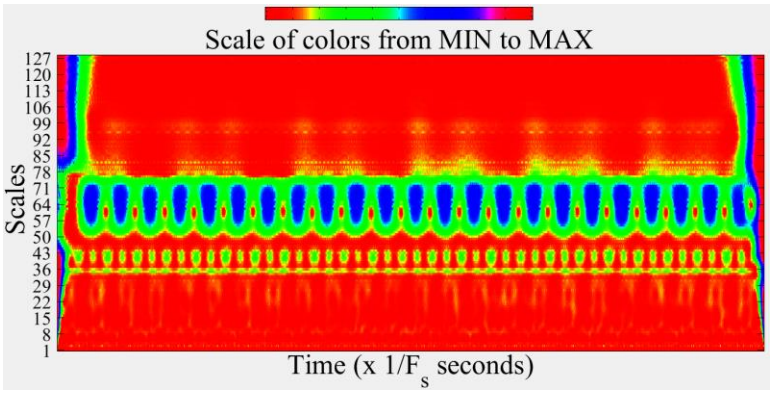
complex CWT magnitude coefficients for the two motor loadings, 30% and 85%, are shown in Figs. 3.3 and 3.4.

The complex CWT magnitude coefficients at lower scale values correspond to higher frequency components, and at higher scale values correspond to lower frequency components. Higher and lower harmonic components may be observed even when the motor has no BRB faults as there might be speed ripple or induced rotor asymmetry due to rotor ellipticity and shaft misalignment.

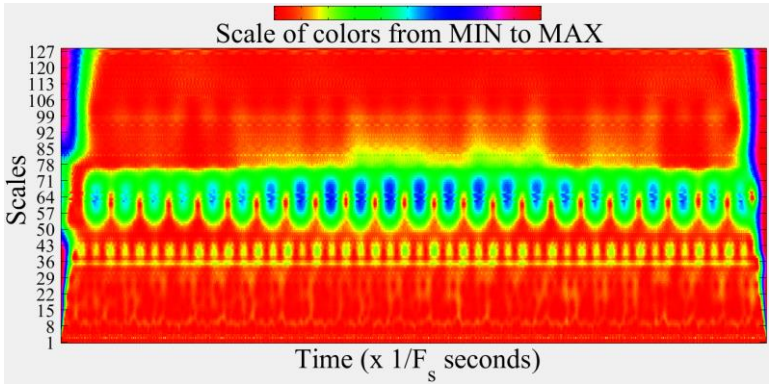
Figs. 3.3(a) and 3.4(a) show the time-scale scalogram representation of the complex CWT magnitude coefficients for healthy motor with 30% and 85% loading, respectively. The coefficients are mostly concentrated within two scale ranges for each graph. For the 30% loading, the lower scale range from 52 to 71, and the upper scale range from 74 to 92; for 85% loading, the lower scale range from 52 to 66, and the upper scale range from 69 to 81. At the areas below or above the lower and upper ranges, very small amount of harmonic components are present. Figs. 3.3(b)-(d) and Figs. 3.4(b)-(d) show the scalograms of complex CWT magnitude coefficients for faulty motor with 1 BRB, 2 BRB, and 3 BRB faults under 30% and 85% loading. Significant harmonic components appear at the three faulty conditions compared to the healthy one. Due to broken rotor bars, the rotor current becomes asymmetrical causing harmonics in upper scale ranges. This further causes ripples in torque and speed and results in harmonics in lower scale ranges. More harmonics appear for the case with the increasing number of BRBs. The severity of the fault can be easily observed.


Scale of colors from MIN to MAX
Scales
127
120
113
106
99
92
85
78
71
64
57
50
43
36
29
22
15
8
1
Time (x 1/F s seconds)

(a)


Scale of colors from MIN to MAX
Scales
127
120
113
106
99
92
85
78
71
64
57
50
43
36
29
22
15
8
1
Time (x 1/F s seconds)

(b)


Scale of colors from MIN to MAX
Scales
127
120
113
106
99
92
85
78
71
64
57
50
43
36
29
22
15
8
1
Time (x 1/F s seconds)

(c)

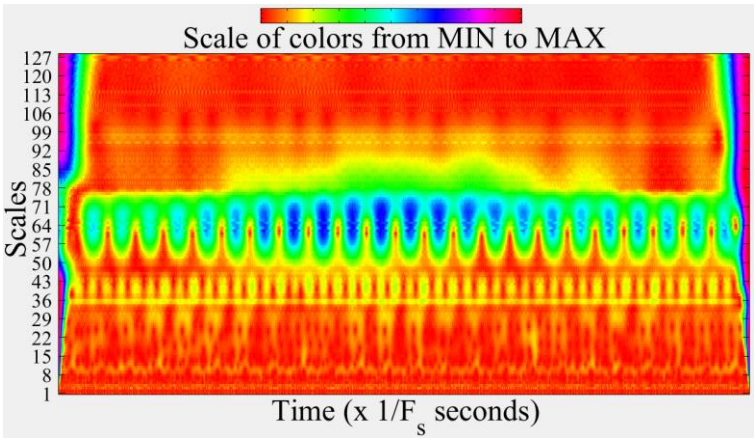
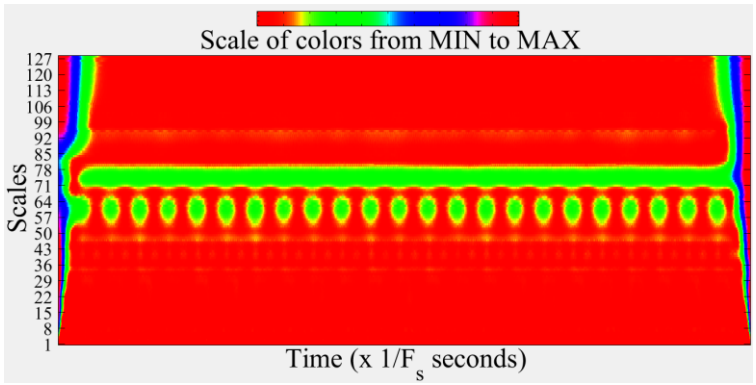
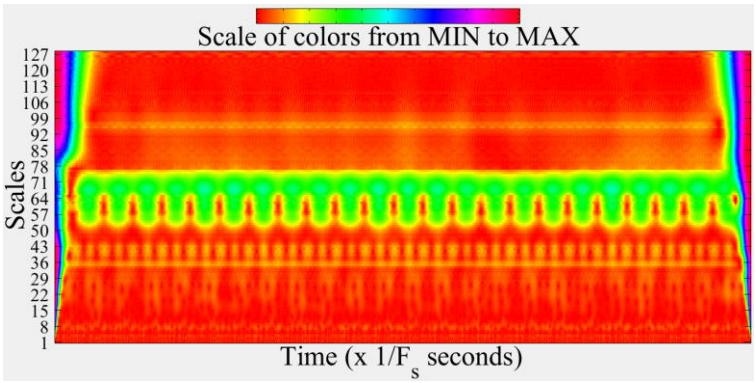


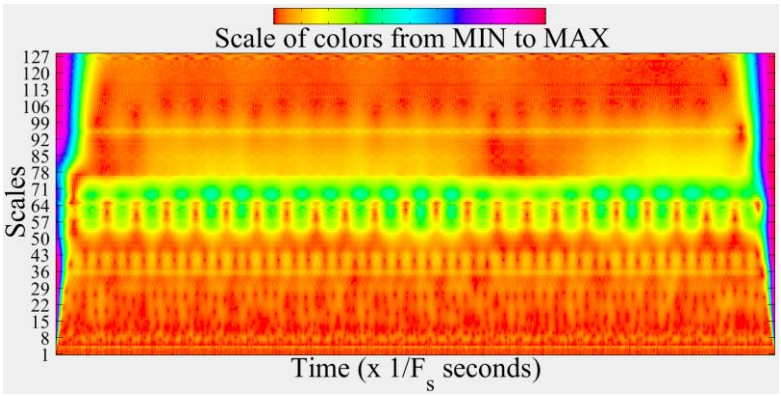
Fig. 3. 3. Scalogram of the complex CWT magnitude coefficients at 30% motor loading: (a) Healthy; (b) 1 BRB; (c) 2 BRB; (d) 3 BRB.



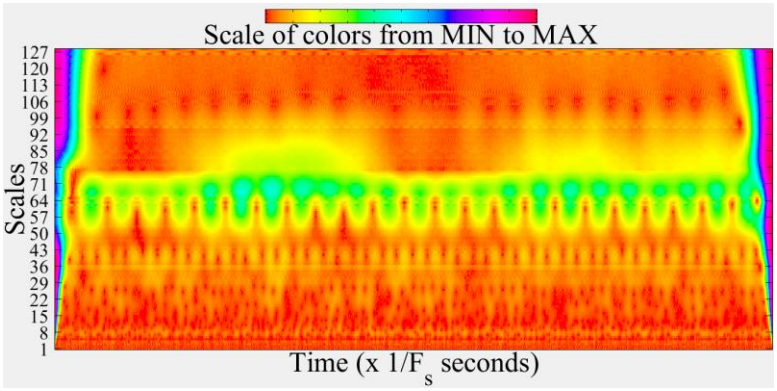
(a)



(b)



(c)



(d)

Fig. 3. 4. Scalogram of complex CWT magnitude coefficients at 85% motor loading: (a) Healthy; (b) 1 BRB; (c) 2 BRB; (d) 3 BRB.

Table 3.2 contains statistical features extracted from the Complex CWT magnitude coefficients. It can be seen that all features increase with the increase of the fault severity except one feature "Mean", which decreases gradually.

Table 3. 2: Statistical Features Extracted from Complex CWT Magnitude Coefficients

Condition	Features	30% Loading	85% Loading
1 BRB	Maximum Value	0.4225	0.7286
	Mean	0.0663	0.0823
	Crest Factor	21.3524	31.2915
	Skewness	17.9150	27.3016
	Kurtosis	479.5	1121.8
2 BRB	Maximum Value	0.4299	0.7376
	Mean	0.0633	0.0789
	Crest Factor	22.8760	32.7231
	Skewness	20.0716	28.2711
	Kurtosis	538.3	1145.9
3 BRB	Maximum Value	0.4310	0.7546
	Mean	0.0631	0.0785
	Crest Factor	24.3992	33.9207
	Skewness	20.9149	29.5825
	Kurtosis	583.8	1228.1

3.4 Conclusion

In this paper, two BRB fault diagnosis algorithms, PSD and 1-D CMW based complex CWT algorithms, for induction motors are implemented using experimental stator current signals measured in a lab. Both methods can successfully identify healthy and faulty conditions of the motor. However, the PSD analysis cannot distinguish the severity of the BRB faults under light loading conditions, although it works well under heavy loading conditions. The complex CWT analysis can successfully distinguish between healthy and faulty motor conditions through significant presence of harmonic components under faulty conditions. The severity of the faults can be observed through the increasing amount of harmonic components in the scalograms of

complex CWT magnitude coefficients. Most statistic features increase with the increase of fault severity. This method works well for any loading conditions. Therefore, the complex CWT method is an effective approach for BRB fault diagnosis for induction motors.

References

[1] A. H. Bonnett and G. C. Soukup, "Cause and analysis of stator and rotor failures in three-phase squirrel-cage induction motors," *IEEE Trans. Ind. Appl.*, vol. 28, no. 4, pp. 921-937, July-Aug. 1992.

[2] A. Gugaliya, G. Singh, and V.N.A. Naikan, "Effective combination of motor fault diagnosis techniques," *2018 International Conference on Power, Instrumentation, Control and Computing (PICC),* pp. 1-5, 2018.

[3] X. Liang, and K. Edomwandekhoe, "Condition monitoring techniques for induction motors," *2017 IEEE Industry Applications Society Annual Meeting,* pp. 1-10, 2017.

[4] J. Cusidó, L. Romeral, J. A. Ortega, J. A. Rosero, and A. G. Espinosa "Fault detection in induction machines using power spectral density in Wavelet decomposition," *IEEE Trans. Ind. Electron.,* vol. 55, no. 2, pp. 633 – 643, Feburary 2008.

[5] Z.K. Peng, and F.L. Chu, "Application of the Wavelet Transform in machine condition monitoring and fault diagnostics: a review with bibliography," *Mechanical Systems and Signal Processing,* vol. 18, no. 2, pp. 199-221, March 2004.

[6] A. Saghafinia, S. Kahourzade, A. Mahmoudi, W. P. Hew, and M. N. Uddin, "On line trained Fuzzy logic and adaptive continuous Wavelet transform based high precision fault detection of IM with broken rotor bars," *2012 IEEE Industry Applications Society Annual Meeting,* pp. 1-8, 2012.

[7] D. Gao, Y. Zhu, X. Wang, K. Yan, and J. Hong, "A fault diagnosis method of rolling bearing based on Complex Morlet CWT and CNN," *2018 Prognostics and System Health Management Conference (PHM Chongqing),* pp. 1101 – 1105, 2018.

[8] H. Li, "Complex Morlet Wavelet amplitude and phase map based bearing fault diagnosis," *2010 8th World Congress on Intelligent Control and Automation,* pp. 6923 – 6926, 2010.

[9] H. Li,"Gear fault diagnosis based on Continuous Morlet Wavelet amplitude and phase map," *2010 7th international Conference on Fuzzy Systems and Knowledge Discovery (FSKD),* vol. 6, pp. 2619 – 2622, 2010.

[10] S. Kay, "Modern Spectral Estimation: Theory and Application," Prentice Hall, Englewood Cliffs, New Jersey, 1998.

[11] S. Kay and S. Marple, ''Spectrum analysis - a modern perspective,'' *Proceedings of the IEEE,* vol. 69, no. 11, pp. 1380-- 1419, November 1981.

[12] F. Filippetti, G. Franceschini, C. Tassoni and P. Vas, "AI techniques in induction machines diagnosis including the speed ripple effect," *IEEE Trans. Ind. Appl.*, vol. 34, no. 1, pp. 98-108, 1998.

[13] B. Yazici, and G. Kliman, "An adaptive statistical time-frequency method for detection of broken bars and bearing faults in motors using stator current," *IEEE Trans. Ind. Appl.*, vol. 35, no. 2, pp. 432-452, 1999.

[14] B. Ayhan, M. Y. Chow, H. J. Trussell, M. H. Song, E. S. Kang, and H. J. Woe, ‘‘Statistical analysis on a case study of load effect on PSD technique for induction motor broken rotor bar fault detection,’’ in *Proc. SDEMPED*, Atlanta, GA, Aug. 24--26, pp. 119--123, 2003.

Chapter 4

Graph-Based Semi-Supervised Learning for Induction Motors Single- and Multi-Fault Diagnosis Using Stator Current Signal

Shafi Md Kawsar Zaman[1], *Student Member*, IEEE, Xiaodong Liang[2], *Senior Member*, IEEE, and Huaguang Zhang[3], *Fellow*, IEEE

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

[2]Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

[3]College of Information Science and Engineering, Northeastern University, Shenyang, Liaoning, China.

A version of this chapter has been published in Proceedings of 2020 Industrial and Commercial Power Systems Technical Conference. Shafi Md Kawsar Zaman developed this work under the supervision of Dr. Xiaodong Liang. Shafi's contributions in this paper are listed as follows:

- Performed literature review for induction motors fault diagnosis.
- Used experimental stator current data to develop a signal processing and machine learning-based fault diagnosis approach.
- Examined results and reported findings.
- Involved writing the paper draft as the first author.

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

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

Abstract- Supervised learning has been commonly used for induction motor fault diagnosis, and requires large amount of labeled samples. However, labeling recorded data is expensive and challenging, while unlabeled samples are available abundantly and contain significant information about motor conditions. In this paper, a graph-based semi-supervised learning (GSSL) approach using both labeled and unlabeled data is proposed. Experimental data for two 0.25 HP induction motors under healthy and faulty conditions are used. Discrete Wavelet Transform (DWT) is employed to extract features from recorded stator current signals. Three GSSL algorithms (local and global consistency (LGC), Gaussian field and harmonic function (GFHF), and greedy-gradient max cut (GGMC)) are evaluated in this study, and GGMC shows superior performance over other two. They are also compared with a supervised learning algorithm, support vector machine (SVM). As induction motors often operate under variable loadings, curve fitting equations are developed based on experimental data to generate training data for untested motor loadings.

Keywords- Graph-based semi-supervised learning (GSSL), greedy-gradient max-cut (GGMC), induction motor, fault diagnosis, Discrete Wavelet Transform (DWT).

4.1 Introduction

Induction motors are the workhorse in various industry sectors due to their compact and robust features with low maintenance costs. Although deemed as reliable, induction motors still fail due to electrical and mechanical faults [1]. Effective induction motor fault diagnosis is essential for critical industrial processes. In the literature, induction motor fault diagnosis can be divided into

three streams [2]-[6]: 1) signature extraction-based approaches, which analyze fault signatures in time- and frequency-domain using recorded current, voltage, power, vibration, temperature, or acoustic signals [1]; 2) model-based approaches, which provide satisfactory performance [4], but require precise motor models for faulty conditions that are often difficult to develop [5]; and 3) knowledge-based approaches, which are based on machine learning techniques, and do not require explicit machine models or load characteristics [3][5][6].

Most knowledge-based approaches use supervised learning (SL) in which a large number of labeled data containing healthy or faulty signatures are required for accurate classification [2]. Artificial neural network (ANN) and hybrid ANN are mostly reported. Hybrid ANN are neural networks combined with the analytical redundancy method [6], Park's Vector pattern learning [7], statistical approach [8], convolutional discriminative feature learning [9], or Fuzzy logic [10]. When used for online fault diagnosis, neural networks are mostly unsupervised [11][12]. Other common SL approaches include support vector machine (SVM), k-nearest neighbor (kNN), ensemble, and decision tree [13][14].

Obtaining datasets in industrial applications may not be difficult as condition data can be continuously monitored, but labeling collected samples requires expert intervention. Also, updating trained machine learning models using new datasets can be complex and computationally expensive [15]. To alleviate this issue, semi-supervised learning (SSL) methods are increasingly used in industrial processes for fault diagnosis [16]-[19]. SSL requires a small number of labeled samples along with a large number of unlabeled samples to construct a classification model [16]. In [16], the data is labeled using an iterative self-labeling process, a differential evolution-based positioning optimization algorithm is used for classification. Several SSL-based fault diagnostic

systems are designed to improve accuracy and efficiency[17][18]. The modified kernel semi-supervised locally linear embedding is used for fault detection of electrical fused magnesia furnace [19].

In the SSL domain, graph-based semi-supervised learning (GSSL) is a promising new paradigm for effective propagation of a limited number of initial labeled data to a large amount of unlabeled data[15],[23]-[28]. Fault detection and classification in PV arrays using GSSL is proposed in [15]. Three SSL algorithms (local and global consistency (LGC), the Gaussian random field (GRF) method, and the graph transduction via alternating minimization (GTAM)) are compared based on simulated and real benchmark datasets in [23]. A bivariate formulation of GSSL that can be solved using a greedy-gradient max cut (GGMC) strategy is proposed in [24]. The extension of GSSL to multi-class classification problems is shown in [25][28]. In [26], a graph construction method over data lying on multiple data manifolds is formulated, which characterizes the global pairwise data similarity using geodesic distance-based joint probability. Residential non-intrusive load monitoring using multi-label GSSL is proposed in [27].

Although GSSL has been applied in various fields using different datasets, to the authors' best knowledge, it hasn't been explored in induction motor fault diagnosis. For the very first time, we propose an GSSL-based induction motor single- and multi-fault diagnosis method using stator current signals in this paper. Three GSSL algorithms (LGC, GFHF, and GGMC) are evaluated, and their performance is compared with a SL algorithm, SVM.

The major contribution of this paper is summarized as follows: 1) An effective GSSL-based approach is proposed for induction motor single- and multi-fault diagnosis using lab/field recorded

stator current signals; and 2) Mathematical equations are developed using experimental data by curve fitting to calculate features for untested motor loadings.

The paper is arranged as follows: the main idea of the proposed approach along with fundamental theory and notations of GSSL are introduced in Section 4.2; experimental setup is shown in Section 4.3; signal processing and feature extraction by DWT is covered in Section 4.4; results analysis is shown in Section 4.5; equations development to calculate features for untested motor loadings are shown in Section 4.6; and conclusions are drawn in Section 4.7.

4.2 Proposed GGMC-Based Fault Diagnosis Approach

The main idea of the proposed approach is demonstrated by the following five steps: 1) conduct experiments for an direct online induction motor under healthy, single- and multi-fault conditions under various motor loadings; 2) record three-phase stator currents using a power quality analyzer for each test case; 3) choose the Discrete Wavelet Transform (DWT) as a signal processing method for feature extraction; 4) conduct classification using the selected GSSL algorithms for fault diagnosis based on experimental data; and 5) develop curve fitting equations to calculate features for untested motor loading conditions. The flowchart of the proposed approach is shown in Fig. 4.1.

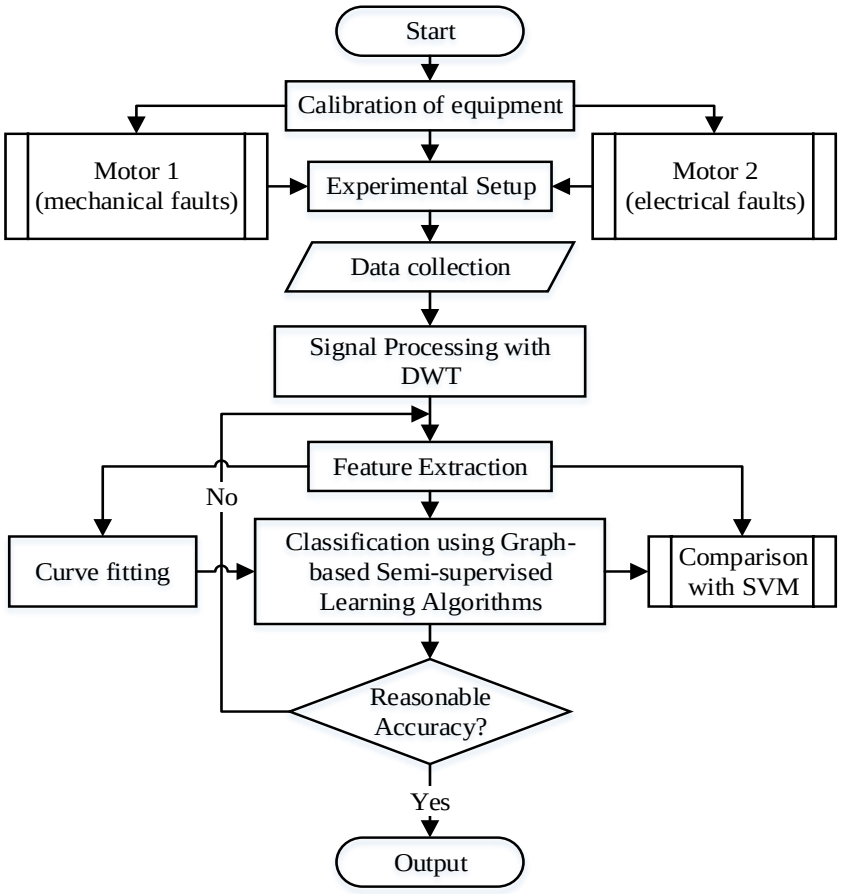


Fig. 4. 1. The flow chart of the proposed approach.

4.2.1 Problem Formulation and Notations for GSSL

It is assumed that the dataset under consideration contains labeled samples $\{(x_1, z_1), \dots, (x_l, z_l)\}$ and unlabeled samples $\{x_{l+1}, \dots, x_{l+u}\}$. The set of labeled inputs are defined as $X_l = \{x_1,\dots,x_l\}$ and the set of unlabeled inputs as $X_u = \{x_{l+1},\dots,x_{l+u}\}$, where l *and* u are the number of labeled and unlabeled inputs, respectively. X_l contains labels $Z_l = \{z_1,\dots,z_l\}$, where, $z_i \in \{1,\dots,c\}$(c represents the number of classes), and $i = 1,2,\dots,l$. GSSL algorithms learn the unknown labels $\{z_{l+1},\dots,z_n\}$ pertaining to the unlabeled data $\{x_{l+1},\cdots,x_n\}$, where typically $l << n$ ($n = l + u$) and estimates a weighted sparse graph G from the input data $X = X_l \cup X_u$. Subsequently, a labeling algorithm uses G and the known labels $Z_l = \{z_1,\dots,z_l\}$ to provide estimate $\hat{Z}_u = \{\hat{Z}_{l+1}, \dots, \hat{Z}_{l+u}\}$, and the actual

labels $Z_u = \{z_{l+1},...,z_{l+u}\}$ are estimated by optimizing an objective function chosen appropriately [33].

In this paper, it is assumed that $G = \{X, E, W\}$ *is* the undirected graph produced from the data X, where the set of vertices is $X = \{x_i\}$, and the set of edges is $E = \{e_{ij}\}$. Each sample x_i is a vertex, and the weight of edge e_{ij} is w_{ij}. Typically, a kernel function $k(\cdot)$ is used over pairs of points to compute weights. The weights for edges are used to build a weight matrix, which is denoted by $W = \{w_{ij}\}$. The vertex degree matrix $D = diag([d_1,...,d_n])$ is defined as $d_i = \sum_{j=1}^{n} w_{ij}$. The graph Laplacian is defined as $\Delta = D - W$ and the normalized graph Laplacian can be represented by

$$L = D^{-1/2}\Delta D^{-1/2} = I - D^{-1/2}WD^{-1/2} \quad (1)$$

The graph Laplacian and its normalized version can be viewed as operators in function space f, which are used to define a smoothness measure in a graph over highly connected regions [29]. The smoothness measurement of functions f using L over a graph is defined by

$$\langle f, Lf\rangle = \sum_i \sum_j w_{ij} \left\| \frac{f(x_i)}{\sqrt{d_i}} - \frac{f(x_j)}{\sqrt{d_j}} \right\|^2 \quad (2)$$

Finally, a label matrix is formulated as $Y = \{y_{ij}\}$ containing label information where $y_{ij} = 1$ if x_i is associated with label j for $j \in \{1,2,...,c\}$(c represents the number of classes), and $y_{ij} = 0$ otherwise. Let $F = f(X)$ be the values of classification function over the data set X. The GSSL methods use W along with the known labels to recover a continuous classification function F by minimizing a predefined objective function on the graph G.

A graph can be built in two typical ways: the ε-neighborhood graph connecting samples within a distance of ε, and the kNN graph connecting k-nearest neighbors. In practice, a kNN graph is a universal approach as it is more adaptive to variation in scale and anomalies in data density, while

an improperly chosen threshold value in the ε-neighborhood graph may result in disconnected components or subgraphs in the dataset or even isolated singleton vertices [30]. In this work, kNN neighborhood graphs are adopted.

4.2.2 Graph Edge Reweighting

After construction of the graph, sparsification is carried out to improve efficiency, accuracy, and robustness to noise during label inference. Graph sparsification removes edges by finding a binary matrix $B \in B\ \{1,0\}^{n\times n}$, where $B_{ij} = 1$ indicates that an edge is present between nodes x_i and x_j, while $B_{ij} = 0$ suggests that the edge is absent. Two possible schemes, binary weighting, fixed Gaussian kernel weighting [23], are commonly considered for graph edge reweighting.

4.2.2.1 Binary Edge Weighting

The most straightforward approach to build the weighted graph is the binary weighting, where the weight 1 is assigned to all linked edges and the weight 0 is assigned to disconnected vertices in the graph. However, this uniform weight on graph edges can be sensitive, especially if the sparsification procedure incorrectly connects some graph vertices.

4.2.2.2 Fixed Gaussian Kernel Edge Weighting

Gaussian kernel weighting is an alternative approach, often used to modulate sample similarity. The edge weight between two connected samples x_i and x_j can be calculated by

$$w_{ij} = B_{ij}\left(-\frac{d^2(x_i,x_j)}{2\sigma^2}\right) \quad (3)$$

Where, the function $d\ (x_i,x_j)$ evaluates the dissimilarity of samples x_i and x_j, and σ is the kernel bandwidth parameter.

4.2.3 Univariate Graph Regularization

In this work, GFHF [31] and LGC [32] methods are considered under the univariate formulation of graph regularization. In LGC and GFHF, an objective function Q is defined, which involves the combined contribution of two penalty terms: the global smoothness Q_{smooth} and local fitting accuracy Q_{fit}. The final prediction function F is obtained by minimizing the objective function by

$$F^* = arg \min_{F \in R^{n \times c}} Q(F) = arg \min_{F \in R^{n \times c}} (Q_{smooth}(F) + Q_{fit}(F)) \quad (4)$$

The formulation of the above objective function is given in [33] for LGC as

$$Q^{LGC}(F) = \|F\|_G^2 + \frac{\mu}{2}\|F - Y\|^2 \quad (5)$$

Where, $\|F\|_G^2$ is function smoothness over the graph G, and $\|F - Y\|^2$ measures the empirical loss of given labeled samples. The function smoothness in LGC is defined by

$$Q_{smooth} = \|F\|_G^2 = \frac{1}{2}\langle F, LF\rangle = \frac{1}{2} tr(F^T L F) \quad (6)$$

The coefficient μ in (5) balances global smoothness and local fitting terms. If $\mu = \infty$ and a standard graph Laplacian quantity Δ for the smoothness term is used, the above framework reduces to the GFHF formulation [31], i.e., the objective function only preserves the smoothness term in GFHF by

$$Q^{GFHF}(F) = tr(F^T \Delta F) \quad (7)$$

4.2.4 Bivariate Graph Regularization and Label Propagation by Greedy Gradient Max-Cut

The optimization problem in LGC and GFHF can be divided into separate parallel problems as the objective function decomposes into additive terms that only depend on individual columns of the prediction matrix [34]. Such decomposition can result in biases if input labels are imbalanced

in proportion. When the graph contains noise and non-separable class manifolds, LGC and GFHF fail to output consistent classification results. To solve this problem, a bivariate optimization framework that explicitly optimizes over both the classification function *F* and the label matrix *Y* is proposed in [24] as follows:

$$Q(F,Y) = \frac{1}{2} tr(F^T LF + \mu(F-Y)^T(F-Y)) \tag{8}$$

Rewriting the objective function as a summation [32] results in a more natural formulation as follows:

$$Q(F,Y) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\left\|\frac{F_i}{\sqrt{d_i}} - \frac{F_j}{\sqrt{d_j}}\right\|^2 + \frac{\mu}{2}\sum_{i=1}^{n}\|F_i - Y_i\|^2 \tag{9}$$

In order to reduce computational complexity, the bivariate optimization problem is reduced to a univariate one [24] for implementing label propagation by GGMC.

The greedy gradient-based strategy is to find local optima by assigning each unlabeled vertex to the label set with minimal connectivity, and iteratively maximizing cross-set edge weights. To alleviate issues, such as biased partitioning, poor graph-cuts and outlier effects, the weighted connectivity between all unlabeled vertices to existing labeled sets is defined to reduce label imbalance across different classes. If the unlabeled vertex x_{i^*} has the least connectivity with label set S_{j^*}, S_{j^*} is updated by adding the vertex x_{i^*} as one greedy step so that the cross-set edge weights are maximized. This greedy search is repeated until all unlabeled vertices are assigned to labeled sets. The weighted connectivity of all unlabeled vertices to labeled sets is recalculated in each iteration of the greedy cut process. The algorithm is known as GGMC, where unlabeled vertices are assigned to labeled sets in a way that lowers the value of objective function *Q* along the steepest descent direction in the greedy step [24][25]. The above formulations are extended to multi-class classification in this work.

4.3 Experimental Setup

In this paper, the proposed approach is developed using lab experimental data. Two identical 4-pole, 0.25-HP, 208–230/460 V, 1725 rpm rated squirrel-cage induction motors (Model LEESON-101649) are tested under various healthy (H), single- and multi-fault conditions. The motors named “Motor 1” and “Motor 2” are treated as sister units.

The five faults applied to Motor 1 are mainly mechanical faults including: 1) an unbalance shaft rotation (UNB); 2) a bearing fault (BF); 3) a multi-fault with BF and UNB (BF+UNB); 4) a multi-fault with BF and one broken rotor bar (BRB) (BF+1BRB); and 5) a multi-fault with BF, UNB, and unbalanced voltage (UV) from the three-phase power supply (BF+UNB+UV)). The five faults applied to Motor 2 are mainly electrical faults including: 1) a UV from the three-phase power supply; 2) one BRB fault (1BRB); 3) two BRBs fault (2BRB); 4) three BRBs fault (3BRB), and 5) a multi-fault with UV and 3BRB (UV+3BRB). Various tests applied to the two motors are shown in Fig. 4.2.

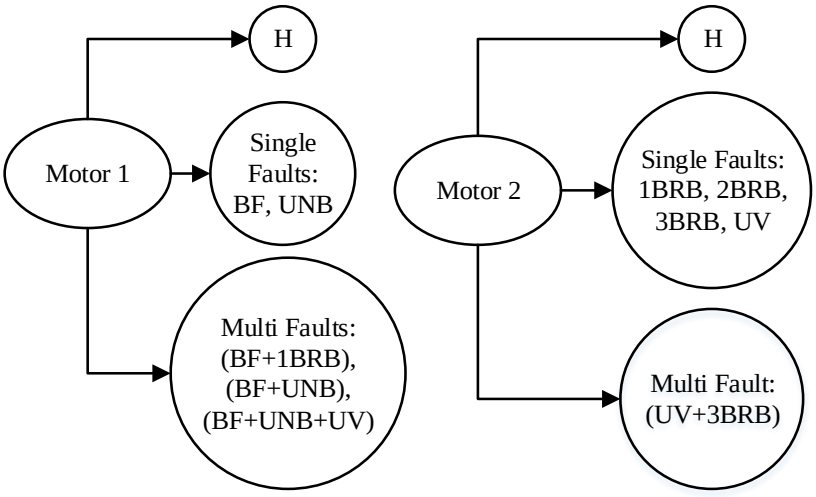


Fig. 4. 2. Two motors with single- and multi-faults.

The experimental test bench is shown in Fig. 4.3, it consists of the direct online three-phase induction motor, and a dynamometer coupled to the motor shaft through a belt pulley serving as the load. The motor loadings are changed by adjusting the dynamometer's control knob. An eight-channel power quality analyzer (PQPro by CANDURA Instruments) is used to record three-phase stator currents. In each test, three-phase stator currents (I_1, I_2, and I_3) are recorded for 2 min with a sampling frequency of 15.38 kHz. A fault creates unbalance inside the motor, and is reflected in the stator current signal.

A BRB fault was created 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. 4.4a); two and three holes with 90° separations were drilled for two and three BRBs faults, respectively (Figs. 4.4b and 4.4c). A general roughness type bearing fault (BF) was realized by the sandblasting process; the outer and inner raceway of the bearing became very rough, as shown in Fig. 4.4d. The UNB is due to uneven mechanical load distribution causing unbalanced shaft rotation, and it was created by adding extra weight on part of the pulley (Fig. 4e). A UV condition was produced by adding extra resistance at the second phase of the power supply. Six motor loadings (10%, 30%, 50%, 70%, 85%, and 100% of full load) were tested for each healthy or faulty condition.

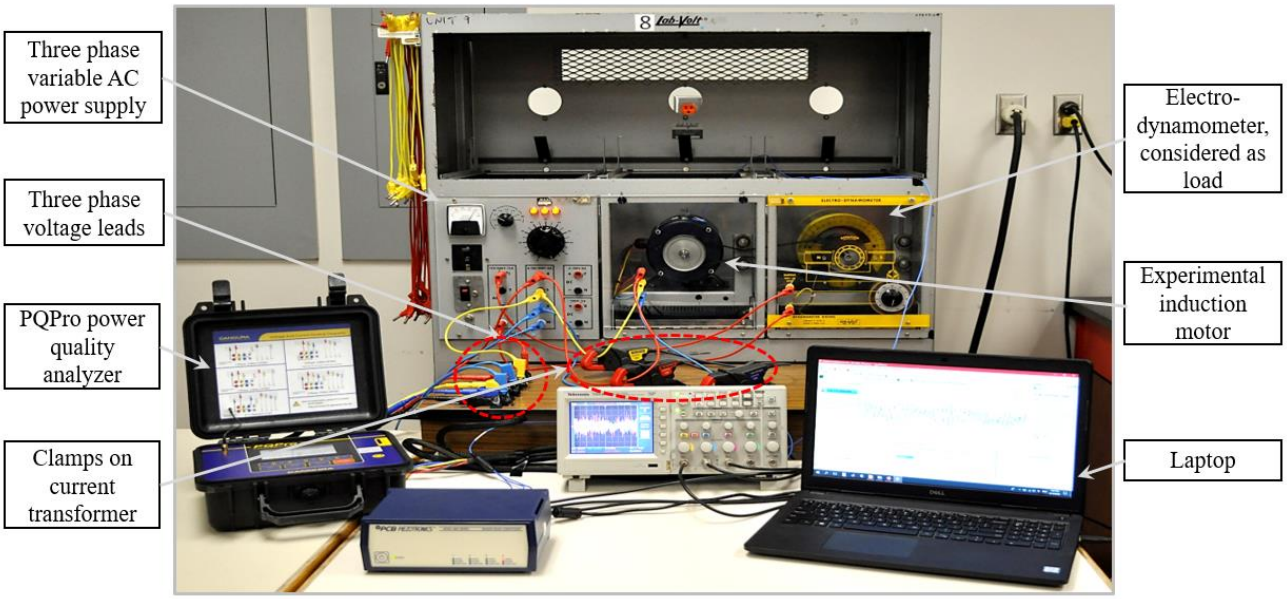


Fig. 4. 3. Experimental test bench.

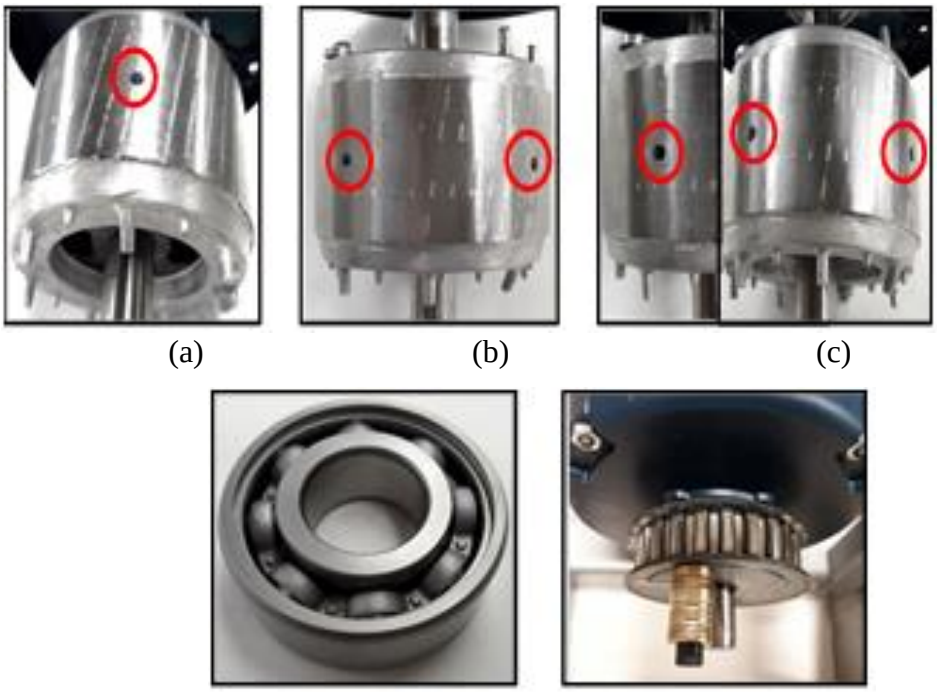


Fig. 4. 4. Different faults implemented on the motors in the lab: (a) 1BRB; (b) 2BRB; (c) 3BRB; (d) BF; and (e) UNB.

4.4 Feature Extraction Using DWT

A wavelet transform is defined as the decomposition of a stationary or nonstationary signal into a set of basic functions consisting of contractions, expansions, and translations of a mother function $\psi(t)$, called the wavelet. The discrete wavelet transform (DWT) [35],[36] uses orthogonal wavelets such as Daubechies wavelet series to decompose a signal into different frequency bands [37]. In this paper, one phase of the measured three-phase stator current I_2 is processed by DWT through MATLAB Wavelet Analyzer toolbox to extract fault features. The length of the dataset was selected uniformly with 90,000 data points for each test, the dataset was further segmented into 10 data windows, each contains 9,000 data points.

In this study, the Daubechies wavelet with four vanishing moments as db4 is selected as the mother wavelet with up to 6^{th} level of decomposition. Ten time-domain statistical features are extracted from the dataset: 1) the maximum value of the data window, 2) the minimum value of the data window, 3) mean, 4) median, 5) median absolute deviation, 6) mean absolute deviation, 7) L1 norm, 8) L2 norm, 9) maximum norm, and 10) standard deviation [22][38].

Motors 1 and 2 each has one healthy and five faulty cases, and thus, there are six class labels for each motor within one data window. Since there are 10 data windows for a dataset under each motor loading, it leads to a total of 6×10=60 class labels for a dataset. Using DWT, 10 features are extracted for each data window under a specific class label, so we have 6×10×10=600 features for each motor under one loading.

Fig. 4.5 shows the DWT processed experimental stator current signal I_2 for Motor 2 under a 1BRB fault and 100% motor loading. Table I shows the corresponding sample of features for the same case. Every set of ten features under a specific class label, such as s_1 at the first row of Table 4.1, is processed using DWT by choosing a data window containing 9,000 sample data points. The remaining nine sets of features (from s_2 to s_{10}) are determined in the same way based on sample data points from nine remaining data windows.

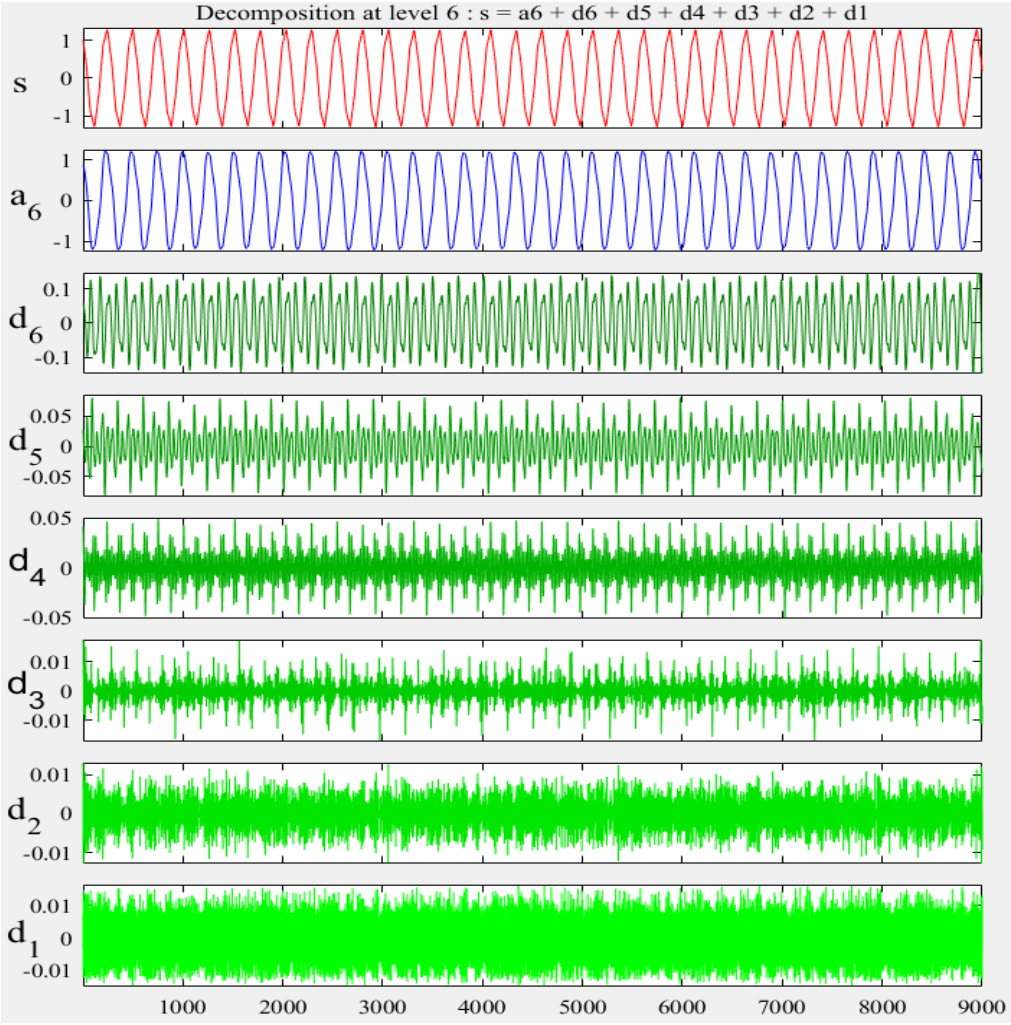


Fig. 4. 5. DWT processed stator current for Motor 2 (1BRB @ 100% loading).

Table 4. 1: Sample of Features extracted by DWT Using Stator Current I_2 (Motor 2, 1 BRB, 100% loading)

Features	Maximum value	Minimum value	Mean	Median	Median Absolute Deviation	Mean Absolute Deviation	L1 norm	L2 norm	Max norm	Standard Deviation
s_1	1.308	−1.309	0.00396	0.01295	0.8438	0.7641	6877	80.47	1.309	0.8482
s_2	1.306	−1.314	−0.00111	−0.00405	0.8440	0.7644	6880	80.53	1.314	0.8489
s_3	1.306	−1.313	−0.00370	−0.00486	0.8486	0.7670	6904	80.74	1.313	0.8511
s_4	1.311	−1.306	−0.00229	−0.00297	0.8467	0.7662	6896	80.67	1.311	0.8503
s_5	1.315	−1.307	0.00202	0.00593	0.8421	0.7627	6865	80.38	1.315	0.8473
s_6	1.296	−1.306	0.00601	0.01052	0.8489	0.765	6885	80.53	1.306	0.8488
s_7	1.300	−1.303	0.00524	0.01079	0.8459	0.7642	6878	80.44	1.303	0.8479
s_8	1.300	−1.304	0.00127	0.00351	0.8408	0.7611	6849	80.21	1.304	0.8456
s_9	1.298	−1.310	−0.00346	−0.00567	0.8432	0.7636	6873	80.39	1.310	0.8474
s_{10}	1.305	−1.314	−0.00365	−0.00432	0.8451	0.7655	6890	80.58	1.314	0.8494

4.5 Result Analysis and Discussion

In this section, fault classification using experimental datasets of the two motors is conducted using the three GSSL algorithms, LGC, GFHF, and GGMC. Their accuracies are evaluated, and compared with that of the SL algorithm, SVM.

For LGC and GGMC, the value of hyper-parameter $\mu = 0.01$ was used for all cases. The three algorithms were all run using 100 independent folds with random sampling to determine the average classification accuracy, and they appeared to require similar run-times to output a prediction. The same graph construction procedure mentioned in Section II is used. The sparsification is performed using the standard approach, k-nearest-neighbors (kNN), and k=4 was used uniformly across all three algorithms. Both binary and fixed Gaussian kernel weighting are used.

4.5.1 Support Vector Machine

SVM is a well-known supervised learning method, and suitable for a dataset where separable and non-separable data manifolds are present. It classifies a dataset into positive and negative classes. A statistical learning theory-based algorithm, known as a support vector, is used to train the dataset. It provides information about the classification and builds the hyperplane. The hyperplane maximizes the separation margin between positive and negative classes [39] and is used to distinguish data points. Different kernel functions are employed in SVM when a nonlinear transformation is required. A kernel function converts nonlinearly separable objects into linearly separable ones, by mapping them into higher dimensional feature space [40]. Common types of kernel functions for SVM include linear, polynomial, and Gaussian radial basis function (RBF) kernels [41].

In this paper, six SVM classifiers named linear, quadratic, cubic, fine Gaussian, medium Gaussian, and coarse Gaussian SVM are used through MATLAB Classification Learner Toolbox. Linear SVM using the linear kernel is the simplest SVM by making a simple linear separation between classes. Quadratic and cubic SVMs use quadratic and cubic kernel functions, respectively. Fine Gaussian SVM using the Gaussian kernel makes detailed distinctions between classes, with kernel scale set to $\sqrt{P}/4$, where P is the number of predictors. Medium Gaussian SVM using the Gaussian kernel makes fewer distinctions than a Fine Gaussian SVM, with kernel scale set to $\sqrt{P}$. Coarse Gaussian SVM using the Gaussian kernel makes coarse distinctions between the classes, with kernel scale set to $4\sqrt{P}$.

4.5.2 Comparison among Three GSSL Algorithms and SVM

In this study, among 60 class labels for a dataset of each motor, only 30 labels (5 labels from each of the six classes to prevent class imbalance) are made available to SVM. When a binary classifier like SVM is used for multi-class classification, the One-vs.-All (OvA) strategy performs better [42], and thus, is adopted for training SVM classifiers in this paper. The performance is validated by the rest 50% of the dataset with unknown labels. In SVM training, five-fold cross-validation is used to prevent overfitting.

The three GSSL algorithms, LGC, GFHF and GGMC, are programed in MATLAB. Unlike SVM, the number of labels for each GSSL algorithm is varied. The three algorithms are designed to initiate from a random stratified selection of 6 known labels to ensure that at least one representative instance of each class is selected. The number of known labels gradually increases from 6 up to 30, as SVM classifiers are also trained for 30 known labels. The two graph edge reweighting schemes, binary edge weighting and fixed Gaussian kernel edge weighting, are implemented for the three algorithms.

The average classification accuracy (by averaging accuracies of 100 iterations) of the three GSSL algorithms and SVM classifiers is tabulated in Tables II and III for the dataset with 30 known and 30 unknown labels. Table II is for the dataset of Motor 1 at 100% loading and Table III is for Motor 2 at 10% loading, with all single- and multi-faults considered.

Tables 4.2 and 4.3 indicate that SVM classifiers all have below 67% classification accuracy, while the three GSSL algorithms perform much better with the accuracy up to 94.7%. GGMC consistently performs better than LGC and GFHF.

Table 4. 2: Comparison Between GSSL and SL Algorithms Based on Classification Accuracies for I_2, Motor 1 (100% Loading)

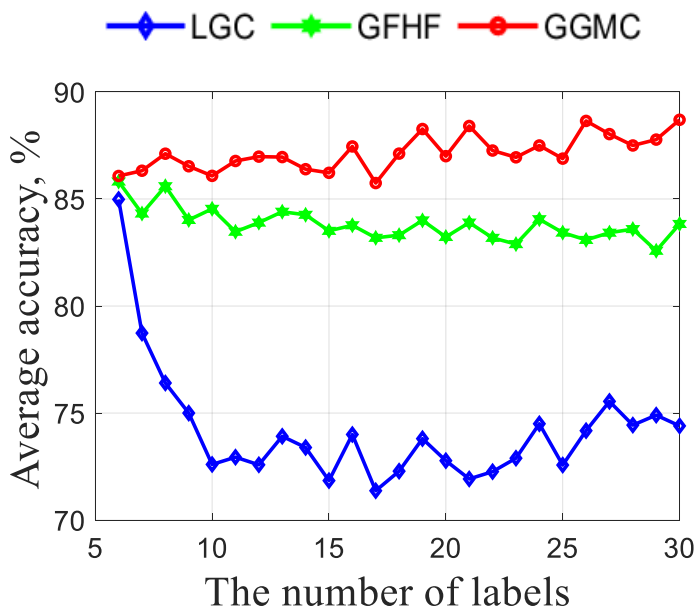
Learning Category	Algorithm	Edge reweighting scheme for GSSL/ Classifier type for SVM	Average classification accuracy in % (30 known and 30 unknown labels)
Graph-based Semi-Supervised Learning (GSSL)	Local and Global Consistency (LGC)	Binary edge weighting	74.4
		Fixed Gaussian kernel edge weighting	78.33
	Gaussian Field and Harmonic Function (GFHF)	Binary edge weighting	83.67
		Fixed Gaussian kernel edge weighting	83.83
	Greedy- Gradient Max- Cut (GGMC)	Binary edge weighting	88.7
		Fixed Gaussian kernel edge weighting	90.03
Supervised Learning (SL)	SVM	Linear SVM	56.7
		Quadratic SVM	66.7
		Cubic SVM	46.7
		Fine Gaussian SVM	40
		Medium Gaussian SVM	43.3
		Coarse Gaussian SVM	33.3

Table 4. 3: Comparison Between GSSL and SL Algorithms Based on Classification Accuracies for I_2, Motor 2 (10% Loading)

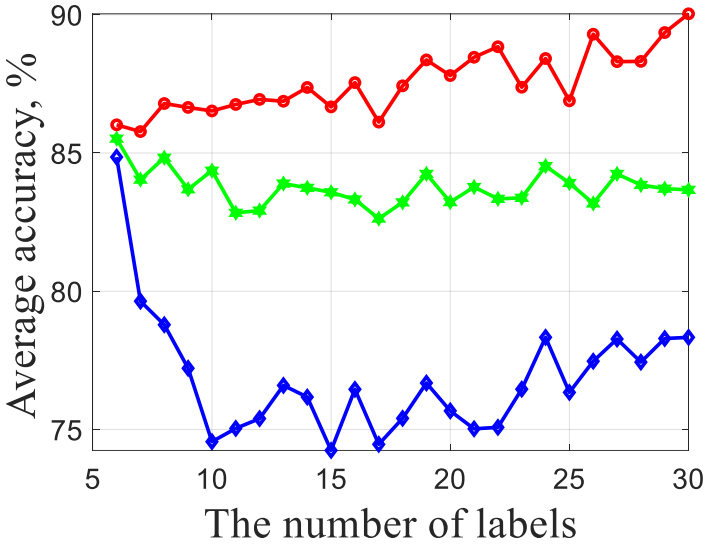
Learning Category	Algorithm	Edge reweighting scheme for GSSL/ Classifier type for SVM	Average classification accuracy in % (30 known and 30 unknown labels)
Graph-based Semi-Supervised Learning (GSSL)	Local and Global Consistency (LGC)	Binary edge weighting	80.7
		Fixed Gaussian kernel edge weighting	79.07
	Gaussian Field and Harmonic Function (GFHF)	Binary edge weighting	91
		Fixed Gaussian kernel edge weighting	84.9
	Greedy- Gradient Max- Cut (GGMC)	Binary edge weighting	94.7
		Fixed Gaussian kernel edge weighting	91.23
Supervised Learning (SL)	SVM	Linear SVM	50
		Quadratic SVM	63.3
		Cubic SVM	56.7
		Fine Gaussian SVM	40
		Medium Gaussian SVM	56.7
		Coarse Gaussian SVM	53.8

In Table 4.2, among the three GSSL algorithms, the lowest average classification accuracy is 74.4% by LGC (with binary edge weighting), and the highest is 90.03% by GGMC (with fixed Gaussian kernel edge weighting). In Table III, the lowest classification accuracy is 79.07% by LGC (with fixed Gaussian kernel edge weighting), and the highest is 94.7% by GGMC (with binary edge weighting). The accuracies achieved by GFHF for both graph edge reweighting schemes are higher than LGC but lower than GGMC. LGC, GFHF, and GGMC all have better classification accuracies using fixed Gaussian kernel edge weighting than binary edge weighting for Motor 1 (mechanical faults), but for Motor 2 (electrical faults), binary edge weighting leads to better accuracies. Nevertheless, GGMC has the best performance.

Figs. 4.6 and 4.7 show the average classification accuracy vs. the number of labels (ranging from 6 to 30 labels) for Motor 1 at 100% loading and Motor 2 at 10% loading, respectively. Binary edge weighting and fixed Gaussian kernel edge weighting are both considered.

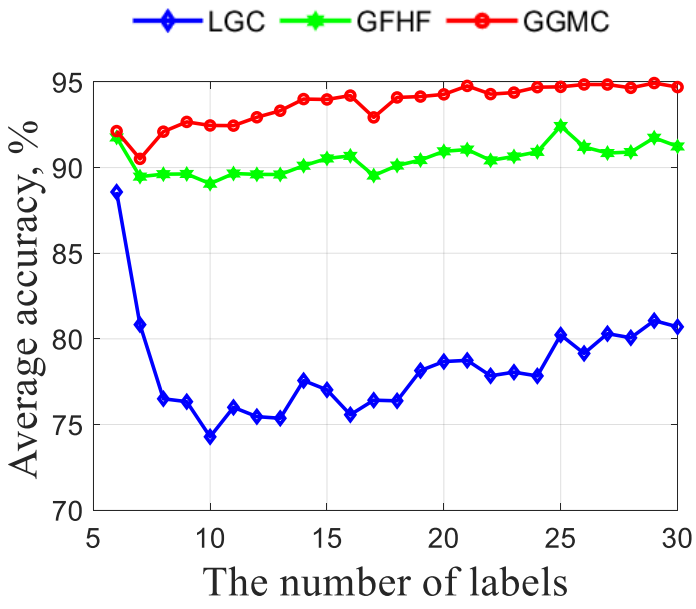


(a) Binary edge weighting

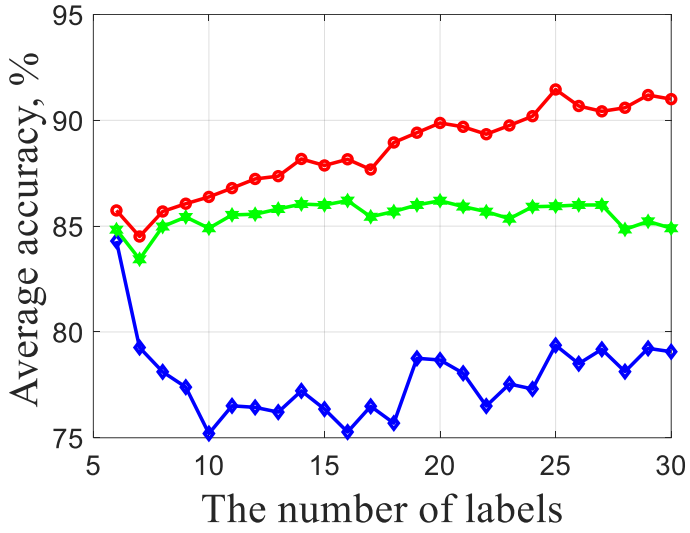


(b) Fixed Gaussian kernel edge weighting

Fig. 4. 6. Average classification accuracy in % vs. the number of labels (for all faults, Motor 1 @ 100% loading).



(a) Binary edge weighting



(b) Fixed Gaussian kernel edge weighting

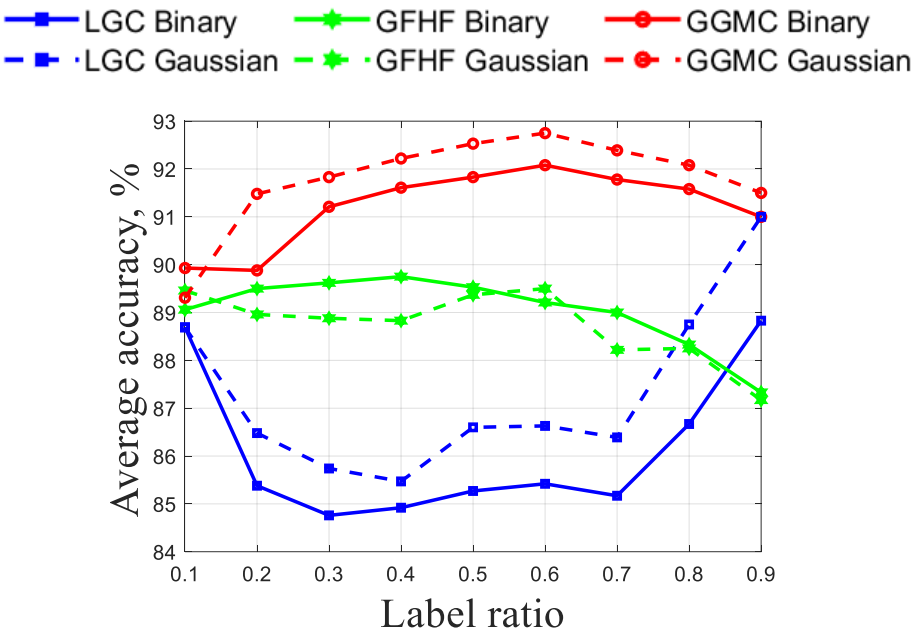
Fig. 4. 7. Average classification accuracy in % vs. number of labels (for all faults, Motor 2 @ 10% loading).

4.5.3 Impact of Label Ratio (LR) on Classification Accuracy

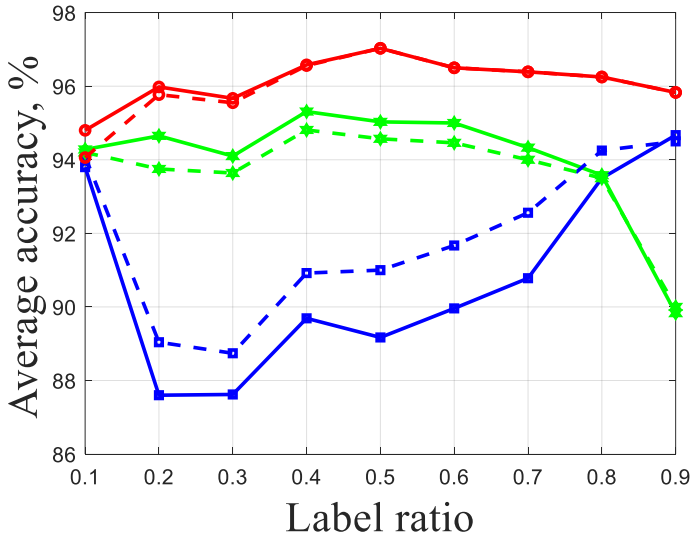
Figs. 4.6 and 4.7 show that the number of labels does affect classification accuracies. The effect of the label ratio (LR) for LGC, GFHF, and GGMC is thus analyzed here. If L is the number of known labels and U is the number of unknown labels, then $LR = L/(L+U)$. Figs. 4.8a and 4.8b show average classification accuracies vs. LR for Motor 1 at 50% loading and Motor 2 at 100% loading, respectively, where LR varies from 0.1 to 0.9. With increasing LR, more labeled data become available, classification accuracies are expected to increase. Fig. 8 shows that GGMC outperforms LGC and GFHF.

In Fig. 4.8a, at LR = 0.6, the average classification accuracy of GGMC reaches the peak of 92.08% with binary edge weighting and the peak of 92.75% with fixed Gaussian kernel edge weighting. Similarly, in Fig. 4.8b, at LR = 0.5, the average accuracy of GGMC reaches the peak of 97.03% for both edge weighting schemes. Larger LR beyond the two peak points will lead to

the performance of GGMC degrades slightly because the superiority of a SSL algorithm starts deteriorating when the percentage of labeled samples reaches a particular value in a dataset [28]. However, average classification accuracies are not that consistent for LGC and GFHF.



(a) Motor 1, 50% loading, for all faults



(b) Motor 2, 100% loading, for all faults

Fig. 4. 8. Average classification accuracy in % vs. label ratio for the motors using binary and fixed Gaussian kernel edge weighting.

Table 4.4 shows a sample of average classification accuracies and standard deviations for LGC, GFHF and GGMC with LR equal to 0.2, 0.3, 0.4 and 0.5 for Motor 1 at 50% loading. The improvement can be observed in terms of standard deviations with increasing LR. GGMC has the least sensitivity towards changes in LR compared to LGC and GFHF, as standard deviations for each LR are the lowest in GGMC, no matter what edge weighting scheme is used. This is also visible in Fig. 4.8. GGMC calculates a normalization of label weight per class, which directly compensates for differences in label proportions [24]. Thus, GGMC is the most stable algorithm among the three. It is observed that standard deviations for GGMC using binary and fixed Gaussian kernels are very close to each another at each LR, so both graph edge reweighting options are excellent for GGMC.

Table 4. 4: Average Classification Accuracies ± Standard Deviations with Respect to Label Ratios for Motor 1 (50% Loading)

Algorithm	Edge weighting scheme	Accuracy (average ± standard deviation), %			
		Label ratio = 0.2	Label ratio = 0.3	Label ratio = 0.4	Label ratio = 0.5
Local and Global Consistency (LGC)	Binary	85.38 ± 1.509	84.76 ± 1.176	84.92 ± 0.9975	85.27 ± 0.8743
	Fixed Gaussian kernel	86.48 ± 1.174	85.74 ± 0.9692	85.47 ± 0.8323	86.6 ± 0.7335
Gaussian Field and Harmonic Function (GFHF)	Binary	89.5 ± 1.016	89.62 ± 0.9987	89.75 ± 0.8537	89.53 ± 0.7596
	Fixed Gaussian kernel	88.96 ± 1.235	88.88 ± 1.213	88.83 ± 1.068	89.37 ± 1.099
Greedy- Gradient Max- Cut (GGMC)	Binary	89.88 ± 0.6039	91.21 ± 0.4779	91.61 ± 0.4662	91.83 ± 0.4461
	Fixed Gaussian kernel	91.48 ± 0.6109	91.83 ± 0.4942	92.22 ± 0.4568	92.53 ± 0.4252

4.6 Features Calculation for Untested Cases

In experiments of this study, the two motors were tested under six loadings: 10%, 30%, 50%, 70%, 85%, and 100%. In real life, the motor may run at any loading factors. Previously, features

for machine learning are extracted by DWT using experimental data, but for untested cases, such as the motor loading is 80%, features cannot be determined as no experimental data available, which results in unavailability of training data for machine learning algorithms. To address this issue, a feature calculation method for untested motor loadings is formulated through curve fitting using experimental data of tested motor loadings.

Bisquare robustness algorithm is used to improve the accuracy of curve fitting equations, where a weight is assigned to each data point based on the distance of that point from the fitted line. The minimization of the weighted sum of squares is carried out. Points that are in proximity of the line get full weight, and points that are farther away from the line get reduced weight. Bisquare is preferable in most applications as it seeks to find a curve so that the bulk of the data is fitted using the conventional least-square approach, and the effect of the outliers is minimized. Bisquare algorithm develops a function by an iterative process, and residuals are calculated by [43]

$$r = \frac{1}{n}\sum_{i=1}^{n-1} w_i \left(f(x_i) - y_i\right)^2 \quad (10)$$

Where, n is the number of data samples, w_i is the i^{th} element of the weights array for data samples, $f(x_i)$ is the fitted model's y-value, and y_i is the i^{th} element of the data set.

4.6.1 Derive Equations for Feature Calculation

To derive curve fitting equations to calculate features for untested motor loadings, the motor loading in percentage is an independent variable, each feature is the dependent variable. *R*-squared values and relative errors between experimental and calculated data using fitted equations are calculated to evaluate the accuracy of the developed equations. R-squared values (ranging from 0 to 1) represent the fitted model's performance following the variance of the actual dataset, and the value close to 1 represents a better fit [44].

Table 4.5 shows the equations to calculate features for Motor 2 with a multi-fault (3BRB + UV) along with R-squared values, where *x* represents the percentage of motor loading, and *y* represents a feature, such as "Mean". The second-order polynomial equations are developed for almost all features except for "mean" and "median". Mean and median use fourth-order polynomial equations. R-squared values are all close to 1 in this table. Relative errors between experimental and calculated data using the fitted equations are shown in Table 4.6 with the highest error equal to 0.7278%.

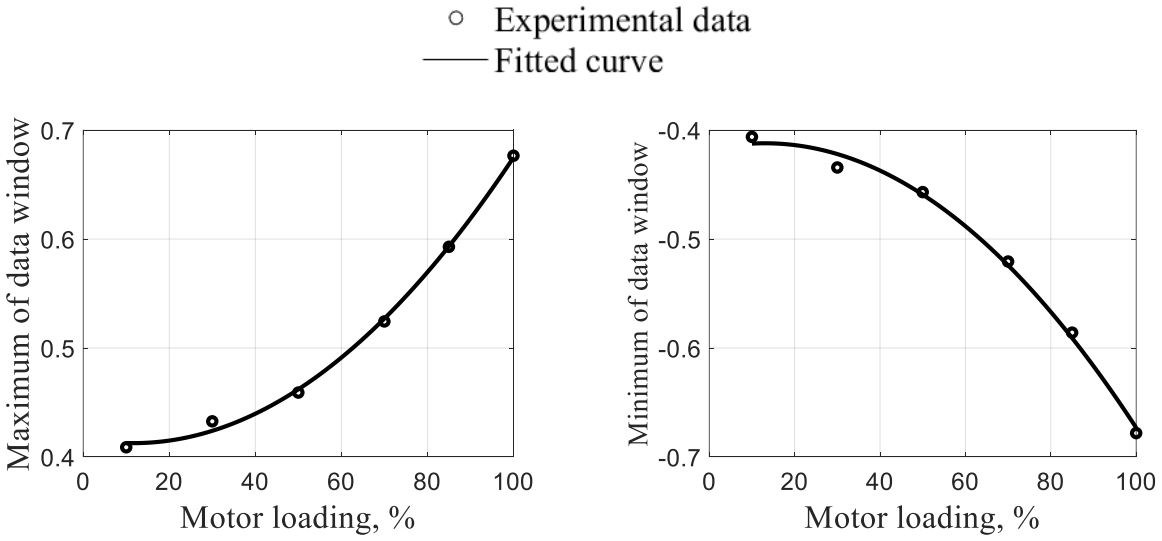
Table 4. 5: Regression Models for Features calculation for Motor 2, a Multi- Fault case (3BRB+UV)

Features Name	Equations	R-square values
Maximum value of data window	$y = 3.359E\text{-}05x^2 - 0.0007837x + 0.4173$	0.9976
Minimum value of data window	$y = -3.455E\text{-}05x^2 + 0.0009034x - 0.418$	0.9944
Mean	$y = -7.95E\text{-}11x^4 + 1.273E\text{-}08x^3 - 4.093E\text{-}07x^2 - 4.707E\text{-}06x + 0.003653$	0.9888
Median	$y = -2.374E\text{-}10x^4 + 1.163E\text{-}07x^3 - 1.469E\text{-}05x^2 + 0.0006294x - 0.001147$	0.9821
Median Absolute Deviation	$y = 1.029E\text{-}05x^2 + 0.00135x + 0.2253$	0.9996
Mean Absolute Deviation	$y = 1.116E\text{-}05x^2 + 0.0007844x + 0.204$	0.9994
L1 Norm	$y = 0.1004x^2 + 7.072x + 1836$	0.9993
L2 Norm	$y = 0.001151x^2 + 0.08647x + 21.48$	0.9994
Maximum Norm	$y = 3.391E\text{-}05x^2 - 0.0008145x + 0.4182$	0.9968
Standard Deviation	$y = 1.213E\text{-}05x^2 + 0.000912x + 0.2264$	0.9994

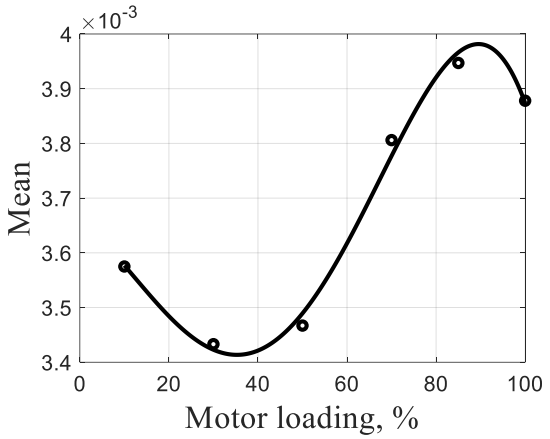
Table 4. 6: Relative Errors Between Experimental and Calculated Data from Fitted Equations for Motor 2, 3BRB+UV, 100% Loading

Features Name	Experiment based data	Calculated data	Relative Error, %
Maximum value of data window	0.6765	0.674823	0.2479
Minimum value of data window	–0.6781	–0.67317	0.7278
Mean	0.003878	0.003874	0.1138
Median	0.007553	0.007515	0.4985
Median Absolute Deviation	0.4645	0.463179	0.2843
Mean Absolute Deviation	0.3955	0.394049	0.3667
L1 Norm	3560	3546.618	0.3759
L2 Norm	41.78	41.64092	0.3329
Maximum Norm	0.6781	0.675852	0.3315
Standard Deviation	0.4404	0.438908	0.3387

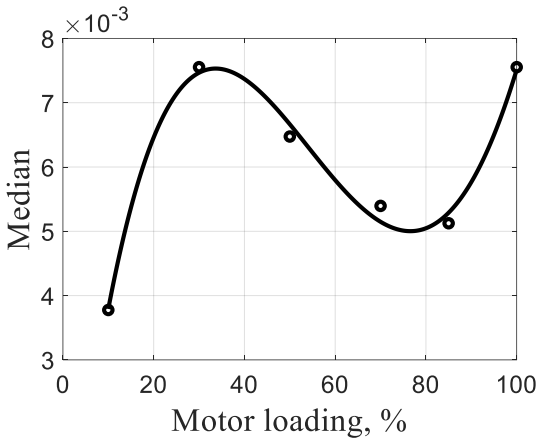
Fig. 4.9 shows the features versus motor loadings for Motor 2 for a multi-fault (3BRB+UV). The dots are DWT processed features using experimental data, while the solid line is determined by fitted equations. Features of other types of faults can be determined in a similar way.



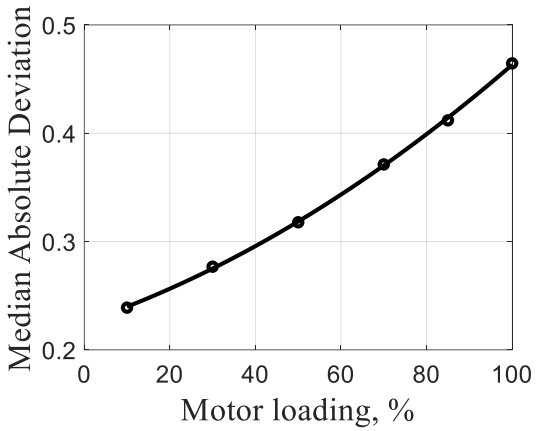
(a) Maximum value of data window (b) Minimum value of data window



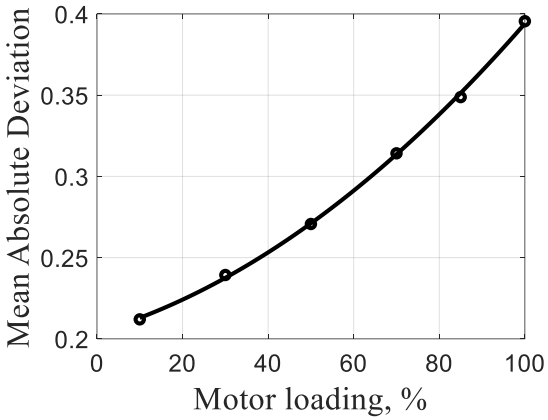
(c) Mean



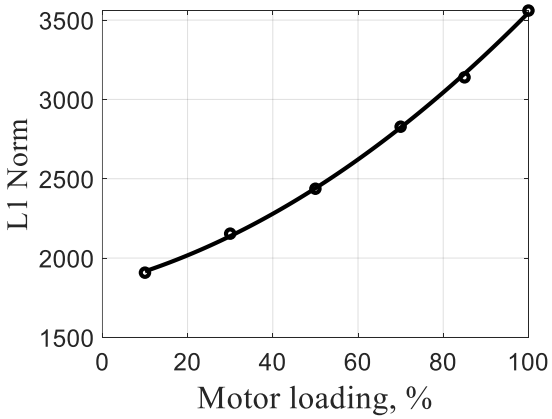
(d) Median



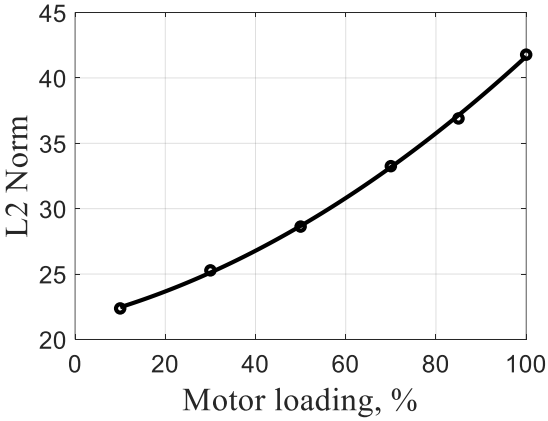
(e) Median absolute deviation



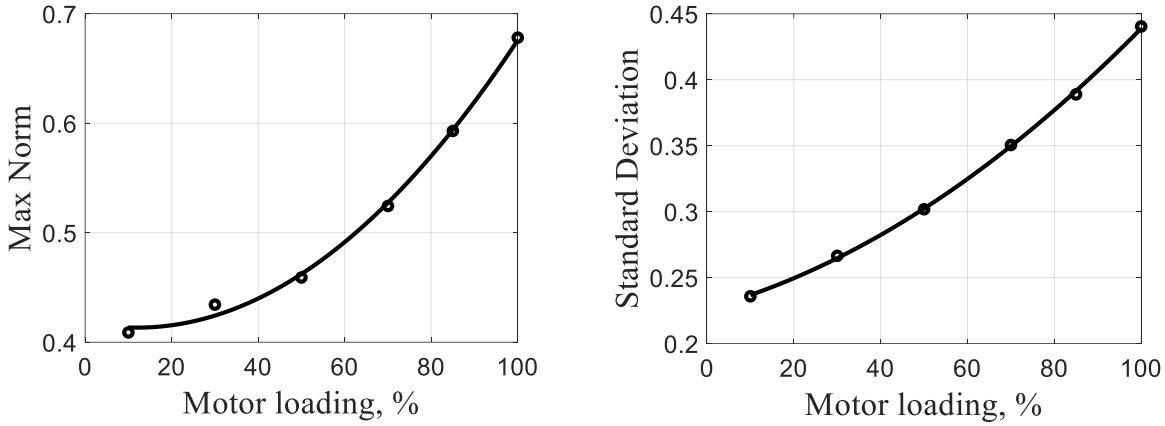
(f) Mean absolute deviation



(g) L1 norm



(h) L2 norm



(i) Max norm (j) Standard deviation

Fig. 4. 9. Features vs. motor loadings through curve fitting technique for Motor 2 with a multi-fault (3BRB+UV).

4.6.2 Machine Learning Results Using Fitted Equations

Feature sets are calculated for two untested loadings for Motor 2, 20%, and 80%, using the developed equations. Fault diagnosis using LGC, GFHF, and GGMC is conducted. Fig. 4.10 shows classification accuracies for all faults. It is found that GGMC outperforms LGC and GFHF for both loadings. With the binary edge weighting, GGMC achieves an average classification accuracy of 89.13% at 20% loading, and 90.63% at 80% loading. With the fixed Gaussian kernel edge weighting, GGMC achieves an average classification accuracy of 88.53% at 20% loading and 90.3% at 80% loading.

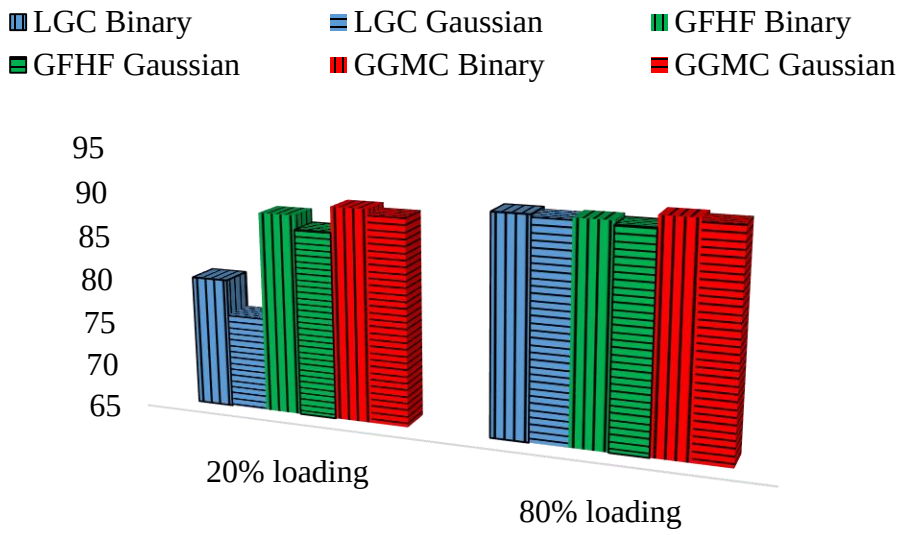


Fig. 4. 10. Classification accuracies for all faults based on I_2 with features extracted from curve fitting equations for Motor 2 at 20% and 80% loadings, using the three GSSL algorithms.

4.7 Conclusion

In this study, an effective GGMC-based induction motor single- and multi-fault diagnosis using the stator current signal is proposed. Three GSSL algorithms, LGC, GFHF, and GGMC, are evaluated, and GGMC shows the best classification accuracy among the three GSSL algorithms. All the GSSL algorithms performed better than the chosen SL method, SVM, for a limited number of labeled data. Therefore, GGMC is proposed in this work and can offer accurate fault diagnosis for various healthy and faulty operating conditions for induction motors.

References

[1] X. Liang and K. Edomwandekhoe, "Condition monitoring techniques for induction motors," *2017 IEEE Industry Applications Society Annu. Meeting*, 1-5 Oct. 2017, pp. 1-10.

[2] X. Dai, and Z. Gao, "From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2226-2238, Nov. 2013.

[3] A. Berzoy, A. A. S. Mohamed and O. Mohammed, "Complex-Vector model of interturn failure in induction machines for fault detection and identification," *IEEE Trans. Ind. Appl.*, vol. 53, no. 3, pp. 2667-2678, May/Jun. 2017.

[4] K. Kim and A. Parlos, "Model-based fault diagnosis of induction motors using non-stationary signal segmentation," *Mechanical Systems and Signal Processing*, vol. 16, no. 2-3, pp. 223-253, 2002.

[5] R. Razavi-Far, E. Hallaji, M. Farajzadeh-Zanjani, and M. Saif, " A semi-supervised diagnostic framework based on the surface estimation of faulty distributions," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1277-1286, Mar. 2019.

[6] H. Su and K. T. Chong, "Induction machine condition monitoring using Neural Network modeling," *IEEE Trans. Ind. Electronics*, vol. 54, no. 1, pp. 241-249, Feb. 2007.

[7] 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. Ind. Appl.*, vol. 36, no. 3, pp. 730-735, May/Jun. 2000.

[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. Ind. Electron.*, vol. 60, no. 9, pp. 4034-4042, Sep. 2013.

[9] W. Sun, R. Zhao, and R. Q. Yan, "Convolutional discriminative feature learning for induction motor fault diagnosis," *IEEE Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1350-1359, Jun. 2017.

[10] S. Altug, Mo-Yuen Chen, and H. J. Trussell, "Fuzzy inference systems implemented on neural architectures for motor fault detection and diagnosis," *IEEE Trans. Ind. Electron.*, vol. 46, no. 6, pp. 1069-1079, Dec. 1999.

[11] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D Convolutional Neural Networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067-7075, Nov. 2016.

[12] M. Seera and C. P. Lim, "Online motor fault detection and diagnosis using a hybrid FMM-CART model," *Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 4, pp. 806-812, Apr. 2014.

[13] 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 Trans. Ind. Appl.*, vol. 55, no. 3, pp. 2378-2391, May/Jun. 2019.

[14] O. Ondel, E. Boutleux, and G. Clerc, "A method to detect broken bars in induction machine using pattern recognition techniques," *IEEE Trans. Ind. Appl.*, vol. 42, no. 4, pp. 916-923, Jul./Aug. 2006.

[15] Y. Zhao, R. Ball, J. Mosesian, J. F. de Palma, and B. Lehman, "Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2848-2858, May 2015.

[16] D. Wu, X. Luo, G. Wang, M. Shang, Y. Yuan, and H. Yan, "A highly accurate framework for self-labeled semisupervised classification in industrial applications," *IEEE Trans. Ind. Informat.*, vol. 14, no. 3, pp. 909-920, Mar. 2018.

[17] M. Grbovic, W. Li, N. A. Subrahmanya, A. K. Usadi, and S. Vucetic, "Cold start approach for data-driven fault detection," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 2264-2273, Nov. 2013.

[18] W. Jiang, Z. Zhang, F. Li, L. Zhang, M. Zhao, and X. Jin, "Joint label consistent dictionary learning and adaptive label prediction for semisupervised machine fault classification," *IEEE Trans. Ind. Informat.*, vol. 12, no. 1, pp. 248-256, Feb. 2016.

[19] Y. Zhang, Y. Fu, Z. Wang, and L. Feng, "Fault detection based on modified kernel semi-supervised locally linear embedding," *IEEE Access*, vol. 6, pp. 479-487, 2018.

[20] R. Razavi-Far, E. Hallaji, M. Farajzadeh-Zanjani, M. Saif, S. H. Kia, and H. Henao, "Information fusion and semi-supervised deep learning scheme for diagnosing gear faults in induction machine systems," *IEEE Trans. Ind. Electronics*, vol. 66, no. 8, pp. 6331-6342, Aug. 2019.

[21] K. Zhang, B. Tang, Y. Qin, and L. Deng, "Fault diagnosis of planetary gearbox using a novel semi-supervised method of multiple association layers networks," *Mechanical Systems and Signal Processing*, vol. 131, pp. 243-260, 2019.

[22] J. Yuan and X. Liu, "Semi-supervised learning and condition fusion for fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 38, no. 2, pp. 615-627, 2013.

[23] T. Jebara, J. Wang, and S. Chang, "Graph construction and b-matching for semi-supervised learning," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009.

[24] J. Wang, T. Jebara, and S. Chang, "Semi-supervised learning using greedy max-cut," *Journal of Machine Learning Research*, vol. 14, pp. 771-800, Mar. 2013.

[25] Z. Zha, T. Mei, J. Wang, Z, Wang, and X. Hua, "Graph-based semi-supervised learning with multiple labels," *Journal of Visual Communication and Image Representation*, vol. 20, no. 2, pp. 97-103, 2009.

[26] M. Fan, X. Zhang, L. Du, L. Chen, and D. Tao, "Semi-supervised learning through label propagation on geodesics," *IEEE Trans. Cybern.*, vol. 48, no. 5, pp. 1486-1499, May 2018.

[27] D. Li and S. Dick, "Residential household non-intrusive load monitoring via graph-based multi-label semi-supervised learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4615-4627, Jul 2019.

[28] S. Wang, X. Guo, Y. Tie, I. Lee, L. Qi, and L. Guan, "Graph-based safe support vector machine for multiple classes," *IEEE Access*, vol. 6, no. 99, pp. 28097-28107, 2018.

[29] S. Butler and F. Chung, "Spectral graph theory," *Handbook of Linear Algebra*, pp. 47, 2006.

[30] M. Maier, U. V. Luxburg, and M. Hein, "Influence of graph construction on graph-based clustering measures," in *Advances in Neural Information Processing Systems*, 2009.

[31] X. Zhu, Z. Ghahramani and J. D. Lafferty, "Semi-supervised learning using gaussian fields and harmonic functions," in *Proc. 20th Int. Conf. Mach. Learn. (ICML-03)*, 2003.

[32] D. Zhou, O. Bousquet, T. Lal, J. Weston, and B. Schölkopf, "Learning with local and global consistency," in *Advances in Neural Information Processing Systems*, 2004.

[33] W. Liu, J. Wang, and S. Chang, "Robust and scalable graph-based semisupervised learning," *Proc IEEE*, vol. 100, no. 9, pp. 2624-2638, 2012.

[34] J. Wang, S.-F. Chang, X. Zhou, and T.C.S. Wong, "Active microscopic cellular image annotation by superposable graph transduction with imbalanced labels," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 1-8, 2008.

[35] 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, Oct. 2006.

[36] 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, Nov. 2010.

[37] A. R. 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, Jun. 2006.

[38] J. J. Saucedo-Dorantes, M. Delgado-Prieto, R. A. Osornio-Rios, and R. de Jesus Romero-Troncoso, "Multifault diagnosis method applied to an electric machine based on high-dimensional feature reduction," *IEEE Trans. Ind. Appl.,* vol. 53, no. 3, pp. 3086-3097, May/Jun. 2017.

[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," *Proc. IEEE Int. Instrum. Meas. Technol. Conf.,* 23-26 May 2016, pp. 1-5.

[40] 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 Electron.,* vol. 29, no. 2, pp. 936-945, Feb. 2014.

[41] X. Liang, Y. He, M. Mitolo, and W. Li, "Support vector machine based dynamic load model using synchrophasor data," *2018 IEEE/IAS 54th Industrial and Commercial Power Systems Technical Conf. (I&CPS),* 7-10 May 2018, pp. 1-11.

[42] R. Rifkin and A. Klautau, "In defense of one-vs-all classification," *Journal of Machine Learning Research,* vol. 5, pp. 101-141, Jan. 2004.

[43] D. C. Hoaglin (Editor), F. Mosteller (Editor), J. W. Tukey (Editor), “Exploring data tables, trends, and shapes,” Vol. 101. John Wiley & Sons, 2011.

[44] 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. Ind. Appl.,* vol. 53, no. 5, pp. 4162-4175, Sept./Oct. 2017.

Chapter 5

Induction Motor Fault Diagnosis Using Graph-Based Semi-Supervised Learning

Shafi Md Kawsar Zaman[1], *Student Member*, IEEE, Xiaodong Liang[2], *Senior Member*, IEEE, and Lihong Zhang[1], *Senior Member*, IEEE

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

[2]Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

A version of this chapter has been accepted by 2020 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE). Shafi Md Kawsar Zaman co-authored the paper under the supervision of Dr. Xiaodong Liang. Shafi's contributions in this paper are listed as follows:

- Performed literature review for machine learning-based induction motor fault diagnosis.
- Implemented signal processing with DWT and semi-supervised learning algorithms using both stator current and vibration data.
- Examined results and reported findings.
- Involved writing the paper draft as the first author.

Dr. Xiaodong Liang provided continuous technical guidance, checked the results, reviewed the manuscript, provided suggestions to accomplish the work, and modified the final version of the manuscript. Dr. Lihong Zhang, as the co-supervisor, reviewed the manuscript as well.

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 the Memorial University of Newfoundland.

Abstract- In this paper, a graph-based semi-supervised learning (GSSL) method is proposed for fault diagnosis of direct online induction motors using stator current and vibration signals. A 0.25 HP induction motor under healthy, single- and multi-fault conditions is tested in the lab. Three-phase stator currents and three-dimensional vibration signals of the motor are recorded simultaneously under steady-state operation in each test. Features for machine learning are extracted from the raw experimental stator current and vibration data using the discrete wavelet transform (DWT). Three GSSL algorithms, local and global consistency (LGC), Gaussian field and harmonic function (GFHF), and greedy-gradient max cut (GGMC), are used in the paper. It is found that both stator current and vibration signals perform well for one individual fault diagnosis using GSSL algorithms, but for classification of a combination of five different faults, the stator current outperforms the vibration signal significantly. Among the three GSSL algorithms, GGMC shows better classification results over LGC and GFHF for both stator current and vibration signals.

Keywords- Induction motor, fault diagnosis, discrete wavelet transform (DWT), graph-based semi-supervised learning (GSSL), greedy-gradient max cut (GGMC).

5.1 Introduction

Induction motors are widely used in various industrial sectors. To prevent production downtime of critical industrial processes, reduce operational costs and improve the system reliability, induction motors fault diagnosis for various electrical and mechanical faults is very important.

With the advancement of signal processing techniques and artificial intelligence, induction motor fault diagnosis has attracted renewed interests in the past decade. There are three streams of research reported in the literature in this area: 1) signature extraction-based approaches, 2) model-based approaches, and 3) knowledge-based approaches. The knowledge-based approaches are more prevalent and data-driven machine learning techniques for both online and offline applications are increasingly employed [1]. Among various machine learning methods, semi-supervised learning shows advantages as only a few numbers of labeled samples are required to infer useful information from a vast amount of unlabeled samples, without requiring any expert involvement [2].

The propagation of the limited number of known labels to the remaining large proportion of unlabeled data can be executed using graph-based semi-supervised learning (GSSL), which is a promising new area in the semi-supervised learning field [3]–[8]. GSSL has been implemented in fault detection and classification in PV arrays in [3]. Comparison of three different graph-based semi-supervised learning algorithms, local and global consistency (LGC), Gaussian random field (GRF), and graph transduction via alternating minimization (GTAM) have been carried out on simulated and benchmark datasets in [4]. A greedy-gradient max cut (GGMC)-based bivariate formulation strategy for GSSL is proposed in [5], and extension of this strategy for multi-class problems is shown in [6][8]. In [7], Multi-label GSSL based residential load monitoring is proposed. Very limited research on induction motor fault diagnosis using GSSL has been reported in the literature. In the few published papers using GSSL, only vibration signals are utilized, and only one individual fault versas healthy machine cases are dealt with [1][9][10].

In this paper, three GSSL algorithms, LGC, Gaussian field and harmonic function (GFHF), and greedy-gradient max cut (GGMC), are used for induction motor fault diagnosis. Both stator current and vibration signals measured simultaneously in the lab for a 0.25 HP induction motor under healthy and various faulty conditions are used. The fault classification performance using the three GSSL algorithms with stator current and vibration signals is evaluated.

The paper is organized as follows: an overview of the GSSL algorithms is introduced in Section 5.2; the experimental setup and feature extraction using DWT is provided in Section 5.3; the result analysis is conducted in Section 5.4, and conclusions are drawn in Section 5.5.

5.2 Overview of GSSL Algorithms and Notations

Assuming that the data set under consideration have both labeled and unlabeled data, GSSL algorithms approximate a weighted sparse graph from the total input data, and provide an estimate of unknown labels using the known ones. The actual labels are determined later on by optimizing a fitness function chosen appropriately. The graph can be formulated in two typical ways: the ε-neighborhood graph connecting samples within a distance of ε, and the kNN graph connecting k-nearest neighbors. In practice, a kNN graph is a more conventional approach as it is more robust to scale variation and abnormalities in data density [11]. As a result, the kNN neighborhood graphs are adopted in all cases in this work.

5.2.1 Graph Edge Re-weighting

Two schemes, binary edge weighting and fixed Gaussian kernel weighting [4], are mostly considered for graph edge reweighting. In binary weighting, the weight 1 is given to all linked edges in the graph, and the weight 0 is assigned as the edge weights of disconnected vertices.

However, this uniform weight on graph edges can be sensitive, especially when the sparsification procedure incorrectly links some of the graph vertices. Gaussian kernel weighting is an alternative approach to binary weighting, which is often used for sample similarity regularization. The edge weight between two connected samples x_i and x_j, can be calculated as

$$w_{ij} = B_{ij}\left(-\frac{d^2(x_i,x_j)}{2\sigma^2}\right) \quad (1)$$

Where, B_{ij} (0 or 1) denotes the connectivity between x_i and x_j, the function d (x_i, x_j) evaluates the dissimilarity of samples x_i and x_j, and σ is the kernel bandwidth parameter.

5.2.2 Three GSSL Methods: LGC, GFHF, and GGMC

In LGC and GFHF algorithms, a fitness function Q is defined, which involves the combined contribution of two penalty terms: the global smoothness Q_{smooth} and local fitting accuracy Q_{fit}. The final prediction function F is obtained by minimizing the fitness function as follows:

$$F^* = arg\min_{F\in R^{n\times c}} Q(F) = arg\min_{F\in R^{n\times c}}\left(Q_{smooth}(F) + Q_{fit}(F)\right) \quad (2)$$

Formulation of the above objective function is given in [12] for LGC by

$$Q(F) = \|F\|_G^2 + \frac{\mu}{2}\|F - Y\|^2 \quad (3)$$

Where, the first term $\|F\|_G^2$ represents function smoothness over graph G and $\|F - Y\|^2$ estimates the empirical loss of given labeled samples. The coefficient μ in (3) provides a balance between global smoothness and local fitting terms. If $\mu = \infty$ is set, the above formulation reduces to the Gaussian field and harmonic function (GFHF) [13].

The optimization problem in LGC and GFHF can be broken up into separate problems as additive terms [14]. Such a decomposition can result in biases if the input labels are not proportionally balanced, which in turn can cause inconsistent classification results. A bivariate

formulation that explicitly optimizes over both the classification function *F* and the label matrix *Y* to solve this drawback was proposed in [5] as follows:

$$Q(F,Y) = \frac{1}{2} tr\left(F^T LF + \mu(F-Y)^T(F-Y)\right) \quad (4)$$

Where, L is the normalized graph Laplacian. This formulation is called GGMC since, in the greedy step, the unlabeled vertices are assigned to labeled sets in a way that lowers the value of fitness function *Q* along the steepest descent direction. The above formulations are extended for multi-class classification cases in this work [5][6].

5.3 Experimental Setup and Feature Extraction

In this paper, a 4-pole, 0.25 HP, 208-230/460 V, 1725 r/min rated squirrel-cage induction motor (Model LEESON-101649) connected direct online is used in a lab experiment for healthy (H) and various single- and multi-faults conditions. A dynamometer coupled to the motor shaft through a belt pulley serves as the load. A total of five faults are applied to the motor including: 1) unbalance shaft rotation (UNB) (by adding extra weight on the pulley); 2) bearing fault (BF) (the general roughness type created with sandblasting); 3) a multi-fault by combining BF and UNB; 4) a multi-fault by combining BF and one broken rotor bar (BRB) (one BRB fault is realized by drilling a hole of a 4.2 mm diameter and 18 mm depth in a rotor bar); and 5) a multi-fault by combining BF, UNB, and unbalanced voltage (UV) of the power supply (UV is done by adding extra resistance at the 2nd phase of the power supply).

An eight-channel power quality analyzer is used to record three-phase currents (I_1, I_2, and I_3) with a sampling frequency of 15.38 kHz. A tri-axial accelerometer with a four-channel sensor signal conditioner mounted on top of the motor near the face end is used to record X, Y, and Z-axes vibration signals with a sampling frequency of approximately 1.3 kHz. The stator currents

and vibration signals were measured simultaneously under steady-state operating conditions. Six different motor loadings (10%, 30%, 50%, 70%, 85%, and 100%) were tested for each heathy or faulty case. From the measurement data, only the stator current from the 2nd phase, I2, and Z-axis vibration were used for feature extraction through DWT. The current and vibration data was selected uniformly for each condition with 90,000 data points under each motor loading, which was further partitioned into the fixed window size of 9,000, resulting in 10 data windows.

DWT available in MATLAB Wavelet Analyzer toolbox is implemented to process the data in each window for feature extraction. Among wavelet families, the Daubechies wavelet with four vanishing moments as db4 is chosen as the mother wavelet with up to the 6th level of decomposition. Ten time-domain statistical features (the maximum and minimum values of the data window, mean, median, median absolute deviation, mean absolute deviation, L1 norm, L2 norm, maximum norm, and standard deviation), are used for GSSL in this paper.

Fig. 5.1 shows the DWT processed Z-axis vibration signal under a BF with 10% motor loading, where s denotes the actual signal, a6 and d1-d6 are approximation and detail levels, respectively.

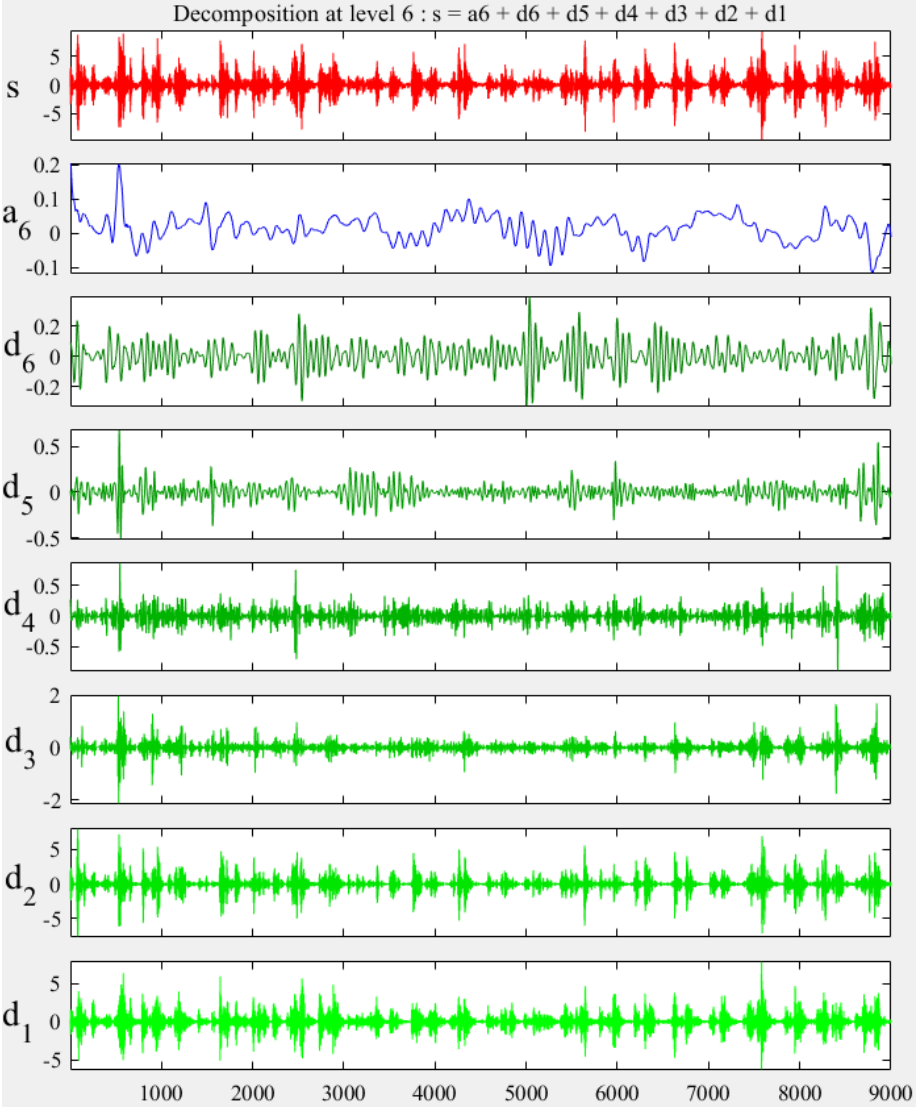


Fig. 5. 1. DWT processed Z-axis vibration signal for the motor with a BF and under 10% motor loading.

5.4 Result Analysis

In this section, the fault classification accuracy for the three GSSL algorithms, LGC, GFHF, and GGMC, are compared using features extracted by DWT from the stator current and vibration signals. For LGC and GGMC, the value of hyper-parameter $\mu = 0.01$ is used across all cases. The three GSSL algorithms implemented in MATLAB all run with 100 independent folds with random sampling using the graph construction procedure mentioned in Section II. It is found that they

require very similar run-time to output a prediction. The sparsification is performed using the k-nearest-neighbors (kNN) approach. For edge weighting, both binary and fixed Gaussian kernel weightings are used in the paper. For the kNN graph construction, k=2 is used uniformly for cases between healthy and one individual fault; while k=4 is used for cases with healthy and all five faults (single- and multi-faults).

5.4.1 GSSL Algorithms for One Individual Fault Classification

For one individual fault diagnosis vs. healthy machine, the GSSL algorithms need a random stratified selection of 2 known labels to ensure at least one representative instance from two different classes is chosen. The number of known labels gradually increases up to 10 in each case, denoting half of the data are labeled. The fault classification performance of the three GSSL algorithms for one individual fault diagnosis is shown in Tables 5.1 and 5.2 using the stator current and vibration signal, respectively. The motor is under 50% loading.

In Table 5.1, using the stator current, the lowest average classification accuracy is 92.5% by LGC and GFHF for a individual multi-fault (BF+1BRB) case vs. a healthy case. The highest average classification accuracy is 100% by GFHF and GGMC for three individual fault cases: a multi-fault (BF+UNB) vs. a healthy case; a multi-fault (BF+UNB+UV) vs. a healthy case; and a single-fault (UNB) vs. a healthy case. Similarly, in Table 5.2, using the vibration signal, the lowest average classification accuracy is 97.1% by LGC with fixed Gaussian kernel edge weighting method for a multi-fault (BF+1BRB) vs. a healthy case. The highest average classification accuracy is 100% by GFHF and GGMC for two individual fault cases: a single-fault (BF) vs. a healthy case; and a multi-fault (BF+UNB+UV) vs. a healthy case. Tables 5.1 and 5.2 indicate that

for individual fault diagnosis, all three GSSL algorithms perform well with comparable accuracies using both stator current and vibration signals.

5.4.2 GSSL Algorithms for All Five Faults Classification

When the classification is performed for all five faults vs. the healthy case, there are six different class labels, the three GSSL algorithms were tuned to commence from a random stratified choice of 6 known labels to ensure one representative instance from six different classes is chosen. In this case, the number of known labels was gradually increased up to 30, denoting half of the data are labeled. Both binary and fixed Gaussian kernel edge weighting schemes are implemented. Table 5.3 shows classification accuracies using the stator current and vibraton signals. Table 5.3 shows very low classification accuracy (below 58%) using the three GSSL algorithms and vibration signal. However, the classification accuracy using the three GSSL algorithms and stator current is good, up to 92.53% for GGMC; in this case, GGMC shows the best performance, and LGC shows the worst.

Fig. 5.2 shows the average fault classification accuracy vs. the number of labels using the three GSSL algorithms with binary and fixed Gaussian kernel edge weighting, respectively. Both the stator current and vibration signals are used for the motor under 50% loading. The average accuracies are computed by averaging classification accuracies of 100 iterations for each GSSL algorithm.

Table 5. 1: Classification accuracies of GSSL algorithms using stator current I_2 and 50% motor loading for one individual fault vs. healthy case

GSSL Algorithm	Edge Reweighting Scheme	Average accuracy, %				
		H vs. BF	H vs. BF+1BRB	H vs. BF+UNB	H vs. BF+UNB+UV	H vs. UNB
Local and Global Consistency (LGC)	Binary edge weighting	97.3	92.5	99.4	99.4	99.4
	Fixed Gaussian Kernel edge weighting	96.3	92.5	99.7	99.8	99.1
Gaussian Field and Harmonic Function (GFHF)	Binary edge weighting	97.3	92.5	100	100	100
	Fixed Gaussian Kernel edge weighting	97.3	92.5	100	100	100
Greedy-Gradient Max-Cut (GGMC)	Binary edge weighting	97.6	94.5	100	100	100
	Fixed Gaussian Kernel edge weighting	97.3	94.5	100	100	100

Table 5. 2: Classification accuracies of GSSL algorithms using Z-axis vibration and 50% motor loading for one individual fault vs. healthy case

GSSL Algorithm	Edge Reweighting Scheme	Average accuracy, %				
		H vs. BF	H vs. BF+1BRB	H vs. BF+UNB	H vs. BF+UNB+UV	H vs. UNB
Local and Global Consistency (LGC)	Binary edge weighting	99.3	97.7	97.8	99.3	97.2
	Fixed Gaussian Kernel edge weighting	98.7	97.1	96.9	98.7	96.6
Gaussian Field and Harmonic Function (GFHF)	Binary edge weighting	100	98.4	97.8	100	97.4
	Fixed Gaussian Kernel edge weighting	100	98.4	97.8	100	97.4
Greedy-Gradient Max-Cut (GGMC)	Binary edge weighting	100	98.4	97.8	100	97.4
	Fixed Gaussian Kernel edge weighting	100	98.4	97.8	100	97.4

Table 5. 3: Classification accuracies of GSSL algorithms using stator current I_2 and Z-axis vibration signal for five faults vs, healthy case

GSSL Algorithm	Edge Reweighting Scheme	Average accuracy %, using Z-axis vibration	Average accuracy %, using the current of phase 2, I_2
Local and Global Consistency (LGC)	Binary edge weighting	55.07	85.27
	Fixed Gaussian Kernel edge weighting	54.3	86.6
Gaussian Field and Harmonic Function (GFHF)	Binary edge weighting	57.63	89.53
	Fixed Gaussian Kernel edge weighting	56.53	89.37
Greedy- Gradient Max- Cut (GGMC)	Binary edge weighting	59	91.83
	Fixed Gaussian Kernel edge weighting	56.6	92.53

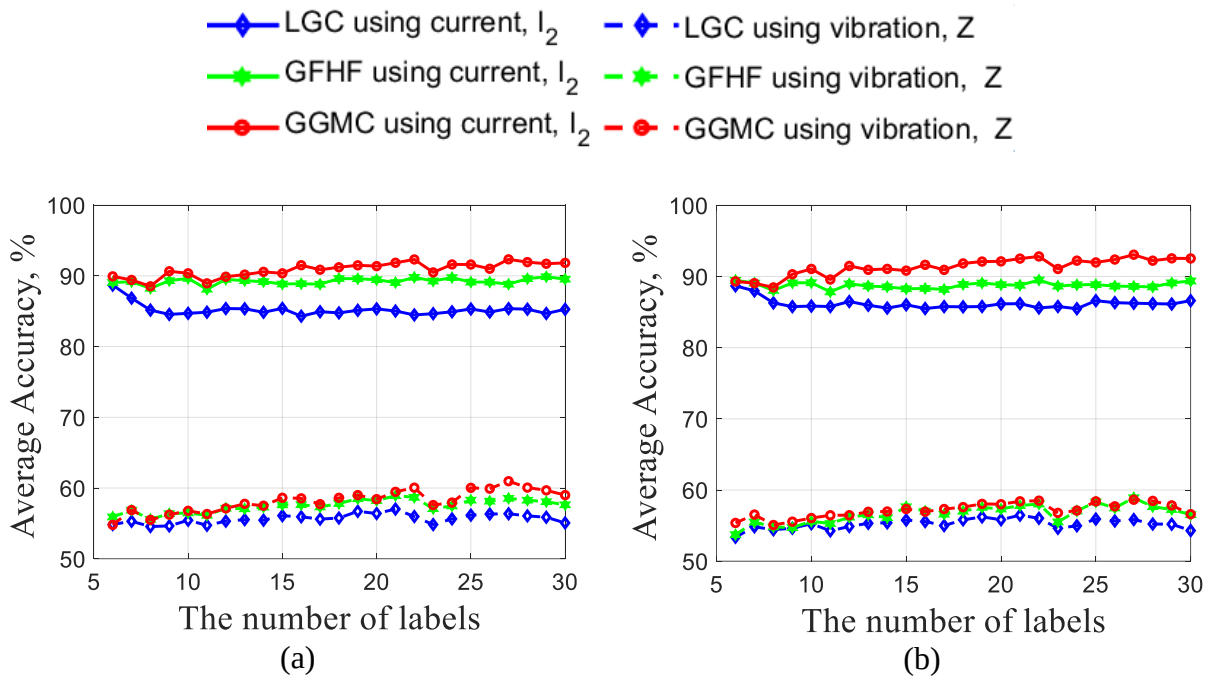


Fig. 5. 2. Average classification accuracy vs. the number of labels for all five faults: (a) Binary edge weighting; (b) Fixed Gaussian kernel edge weighting.

5.5 Conclusion

In this study, induction motor fault diagnosis method is developed using GSSL algorithms and experimental stator current and vibration signals. Three GSSL algorithms, LGC, GFHF and GGMC, are evaluated in the paper. For individual fault diagnosis, all three GSSL algorithms perform well with comparable accuracies using both stator current and vibration signals. For five faults classification, GGMC shows the best performance using the stator current, while none of the three GSSL algorithms perform well using vibration signal. It is recommended that the GGMC algorithm combined with the stator current signal should be used to obtain consistently good fault classification performance for induction motors.

References:

[1] R. Razavi-Far, E. Hallaji, M. Farajzadeh-Zanjani, and M. Saif, "A semi-supervised diagnostic framework based on the surface estimation of faulty distributions," *IEEE Trans. Ind. Inform.*, vol. 15, no. 3, pp. 1277–1286, March 2019.

[2] D. Wu, X. Luo, G. Wang, M. Shang, Y. Yuan, and H. Yan, "A Highly Accurate Framework for Self-Labeled Semisupervised Classification in Industrial Applications," *IEEE Trans. Ind. Inform.*, vol. 14, no. 3, pp. 909–920, March 2018.

[3] Y. Zhao, R. Ball, J. Mosesian, J.-F. de Palma, and B. Lehman, "Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2848–2858, May 2015.

[4] T. Jebara, J. Wang, and S.-F. Chang, "Graph construction and b-matching for semi-supervised learning," *Proceedings of the 26th annual international conference on machine learning*, pp. 441–448, 2009.

[5] J. Wang, T. Jebara, and S.-F. Chang, "Semi-supervised learning using greedy max-cut," *J. Mach. Learn. Res.*, vol. 14, no. Mar, pp. 771–800, 2013.

[6] Z.-J. Zha, T. Mei, J. Wang, Z. Wang, and X.-S. Hua, "Graph-based semi-supervised learning with multiple labels," *J. Vis. Commun. Image Represent.*, vol. 20, no. 2, pp. 97–103, 2009.

[7] D. Li and S. Dick, "Residential household non-intrusive load monitoring via graph-based multi-label semi-supervised learning," *IEEE Trans. Smart Grid*, vo. 10, no. 4, pp. 4615 – 4627, July 2019.

[8] S. Wang, X. Guo, Y. Tie, I. Lee, L. Qi, and L. Guan, "Graph-based safe support vector machine for multiple classes," *IEEE Access*, vol. 6, pp. 28097–28107, May 2018.

[9] W. Jiang, Z. Zhang, F. Li, L. Zhang, M. Zhao, and X. Jin, "Joint label consistent dictionary learning and adaptive label prediction for semisupervised machine fault classification," *IEEE Trans. Ind. Inform.*, vol. 12, no. 1, pp. 248–256, Feb. 2016.

[10] R. Razavi-Far *et al.*, "Information Fusion and Semi-Supervised Deep Learning Scheme for Diagnosing Gear Faults in Induction Machine Systems," *IEEE Trans. Ind. Electron.*, vol. 66, no. 8, pp. 6331–6342, Aug. 2019.

[11] M. Maier, U. V. Luxburg, and M. Hein, "Influence of graph construction on graph-based clustering measures," *Proceedings of the Advances in neural information processing systems*, pp. 1025–1032, 2009.

[12] W. Liu, J. Wang, and S.-F. Chang, “Robust and scalable graph-based semisupervised learning,” *Proc. IEEE*, vol. 100, no. 9, pp. 2624–2638, Sept. 2012.

[13] X. Zhu, Z. Ghahramani, and J. D. Lafferty, “Semi-supervised learning using gaussian fields and harmonic functions,” *Proceedings of the 20th International Conference on Machine Learning*, pp. 912–919, 2003.

[14] J. Wang, S.-F. Chang, X. Zhou, and S. T. Wong, “Active microscopic cellular image annotation by superposable graph transduction with imbalanced labels,” *Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2008.

Chapter 6

Conclusion

6.1 Summary and Conclusions

In this thesis, effective fault diagnosis methods for direct online induction motors using signal processing and machine learning are developed. The main research outcomes of Chapters 3, 4, and 5 are summarized as follows:

In Chapter 3, two BRB fault diagnosis algorithms for direct online induction motors, PSD and 1-D CMW based complex CWT algorithms, are implemented using experimental stator current signals measured in a lab. Both methods can successfully identify the healthy and faulty conditions of the motor. However, the PSD analysis cannot distinguish the severity of BRB faults under light loading conditions. The complex CWT analysis can successfully differentiate between healthy and faulty motor conditions through the significant presence of harmonic components under fault conditions. The severity of faults can be observed through the increasing amount of harmonic components in the scalograms of complex CWT magnitude coefficients. This method works well for any motor loading conditions. Therefore, the complex CWT method is a practical approach for BRB fault diagnosis for direct online induction motors.

In Chapter 4, an effective single- and multi-fault diagnosis method for direct online induction motors is demonstrated using a graph-based semi-supervised learning approach. Three GSSL algorithms- LGC, GFHF, and GGMC, are evaluated based on features extracted by DWT from experimental stator currents, which carry signatures of healthy, single-, and multi-fault motor states. All three GSSL algorithms perform better than the selected supervised learning algorithm

for a given number of labeled data. GGMC is more effective than LGC and GFHF. To enable machine learning for untested motor operating conditions, mathematical equations to calculate features for such untested conditions are developed using curve fitting based on features extracted from experimental data for tested conditions.

In Chapter 5, the induction motor fault diagnosis method is developed using experimental stator current and vibration signals. Three GSSL algorithms, LGC, GFHF, and GGMC, are evaluated in the paper. For individual fault diagnosis, all three GSSL algorithms perform well with comparable accuracies using both stator current and vibration signals. For all faults classification, GGMC shows the best performance using the stator current, while none of the three GSSL algorithms perform well using vibration signals. GGMC, combined with stator current signals, can offer consistently good fault classification performance.

The following conclusions of the thesis can be drawn:

1. CWT is a better signal processing approach compared to PSD estimates for broken rotor bar (BRB) fault detection of direct online induction motors.
2. Among the three GSSL algorithms considered in this thesis for single- and multi-fault diagnosis, GGMC outperforms LGC and GFHF with both stator current and vibration signals.
3. Stator current signal is a more practicable choice for single- and multi- fault diagnosis of direct online induction motors.

6.2 Future Works

- The future research can be extended to develop fault diagnosis methods for variable frequency drive (VFD)-driven induction motors. The feasibility of the proposed methods

such as the CWT and PSD estimates based stator current signature analysis, the graph-based semi-supervised learning based single- and multi-fault diagnosis with both stator current and vibration data can be verified in the future for VFD-fed induction motors.

- Other monitoring signals, such as voltage and instantaneous power, may also be considered in future work. Both signal processing and machine learning methods can be applied to validate the significance of these signals in comparison to stator current and vibration signals, and their performance can also be evaluated.

List of Publications (Since Fall 2018)

Refereed Journal Papers (published or submitted):

[1] Mohammad Zawad Ali, Md Nasmus Sakib Khan Shabbir, **Shafi Md Kawsar Zaman**, and Xiaodong Liang, "Single- and Multi-Fault Diagnosis Using Machine Learning for Variable Frequency Drive-Fed Induction Motors," *IEEE Transactions on Industry Applications*, vol. 56, no. 3, pp. 2324-2337, May/June 2020.

[2] **Shafi Md Kawsar Zaman**, Xiaodong Liang, and Weixing Li, "VFD-Fed Induction Motor Fault Diagnosis Using Wavelet Packet Decomposition and Graph-Based Semi-Supervised Learning," submitted to *IEEE Transactions on Energy Conversion*, pp. 1-8 (under review).

[3] **Shafi Md Kawsar Zaman**, Xiaodong Liang, and Huaguang Zhang, "Graph-Based Semi-Supervised Learning for Induction Motors Single- and Multi-Fault Diagnosis Using Stator Current Signal," submitted to *IEEE Transactions on Industry Applications*, pp. 1-10 (under review).

Refereed Conference Papers (published or accepted):

[4] **Shafi Md Kawsar Zaman**, Xiaodong Liang, and Lihong Zhang, "Induction Motor Fault Diagnosis Using Graph-Based Semi-Supervised Learning," accepted by *33rd Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1-5, London, Ontario, Canada, April 26-29, 2020.

[5] **Shafi Md Kawsar Zaman**, Xiaodong Liang, and Huaguang Zhang, "Graph-Based Semi-Supervised Learning for Induction Motors Single- and Multi-Fault Diagnosis Using Stator Current Signal," *Proceedings of 56th IEEE Industrial and Commercial Power System (I&CPS) Technical Conference*, pp. 1-10, Las Vegas, Nevada, United States, April 27-30, 2020.

[6] **Shafi Md Kawsar Zaman**, Hla U May Marma, and Xiaodong Liang, "Broken Rotor Bar Fault Diagnosis for Induction Motors Using Power Spectral Density and Complex Continuous Wavelet Transform Methods," *Proceedings of IEEE Canadian Conference of Electrical and Computer Engineering (CCECE) 2019*, pp. 1-4, Edmonton, AB, Canada, May 5-8, 2019.

[7] 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", *Proceedings of 54th IEEE Industry*

Applications Society (IAS) Annual Meeting, pp. 1-14, Baltimore, Maryland, United States, September 29th - October 3rd, 2019.

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

[8] **Shafi Md Kawsar Zaman**, and Xiaodong Liang, "Power System Stability Evaluation Using Dynamic Model of a Fixed Speed Induction Generator," *2018 Annual IEEE Newfoundland Electrical and Computer Engineering Conference (NECEC)*, pp. 1-5, St. John's, Canada, November 2018.