



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

DOCTOR OF PHILOSOPHY

DATA-DRIVEN FAULT DIAGNOSIS AND PROGNOSIS
FRAMEWORK FOR BEARINGS USING ADVANCED SIGNAL
PROCESSING AND MACHINE LEARNING TECHNIQUES



The Graduate School
of the University of Ulsan
Department of Computer Engineering

M M MANJURUL ISLAM

DATA-DRIVEN FAULT DIAGNOSIS AND PROGNOSIS
FRAMEWORK FOR BEARINGS USING ADVANCED SIGNAL
PROCESSING AND MACHINE LEARNING TECHNIQUES

DISSERTATION

Submitted to
the Graduate School of the University of Ulsan
In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy

by

M M Manjurul Islam

Department of Computer Engineering
Ulsan, Republic of Korea

June 2019

**Data-Driven Fault Diagnosis and Prognosis Framework for
Bearings Using Advanced Signal Processing and Machine
Learning Techniques**

Supervisor: Professor Jong-Myon Kim

A Dissertation

Submitted to
the Graduate School of Electrical, Electronic, and Computer
Engineering, University of Ulsan
In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
(Computer Engineering)

by

M M Manjurul Islam

University of Ulsan
Ulsan, Republic of Korea
June 2019

Copyright ©2019 – M M Manjurul Islam

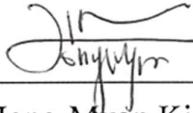
All rights reserved

**Data-Driven Fault Diagnosis and Prognosis Framework for
Bearings Using Advanced Signal Processing and Machine
Learning Techniques**

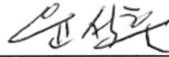
This certifies that the dissertation of M M Manjurul Islam is approved.



Professor Young-Keun Kwon, Committee Chair



Professor Jong-Myon Kim, Advisor



Professor Seok-Hoon Yoon, Committee Member



Professor Cheol-Hong Kim, Committee Member



Professor Ki-Chang Im, Committee Member

Department of Computer Engineering
Ulsan, Republic of Korea

May 2019

VITA

M M Manjurul Islam was born in Homna, Cumilla, Bangladesh, in 1985. He completed his bachelor's in Computer Science and Engineering from Rajshahi University of Engineering and Technology, Bangladesh, in 2007. Since September 2014, he has been working for his Combined Master's and Ph.D. Degree in the Department of Computer Engineering, University of Ulsan, South Korea, under the supervision of Professor Jong-Myon Kim. His Ph.D. dissertation is based on exploring Artificial Intelligence techniques for data-driven Diagnostics and Prognostics in Condition-based Maintenance (CBM). His current research interests include data-driven anomaly detection, diagnostics, prognostics and health management, artificial intelligence, machine learning, signal processing, high-performance computing. He has been a part of several interesting projects related to the diagnostics and prognostic of time-varying, non-linear, non-stationary Systems, embedded systems-, multiprocessor system-on-a-chip (MPSoC)-, GPGPU-based high-performance application development, and vision-based inspection and security.

DEDICATION

My father, Mohammad Nurul Islam, who always wished me for this moment. I miss him every day. He passed away during my journey toward Ph.D. before the final moment of my Ph.D. This thesis is dedicated to the memory of my father.

ACKNOWLEDGMENTS

This has been a long journey with stumbles and rejuvenation, desolation and hope, joy and apprehension, and a sense of achievement. But in the end, it was all worth it. But in this entire expedition, one thing has remained constant—the total guidance by my supervisor Professor Jong-Myon Kim and his timely interventions in the research work to keep it on track. Without his help, the vision of both this research work and the arc of the research direction would be hazy or to be apt short-sighted. I thank Prof. Kim for all his guidance, trust and support in this long journey of my studies. I am also thankful to my Ph.D. supervisory committee for taking out the time to invest in the improvement of my dissertation.

Most importantly, I would like to gratefully thanks the support and well wishes of my dear parents and family members who always love and guide me wholeheartedly in every situation of life. I have to thank my elder brothers who all have been a source of constant support, love, and care to make me go in life. They have focused all their attention and resources on me and understandably all their hopes. I wish I could be able to meet up to their hopes soon. Especially, I would like to express my profound gratitude and appreciation to my loving and supportive wife, Irin, and our adorable kid, Adriyan. My dear ones have always been a steady spiritual support for me during the study and completing this dissertation.

My colleagues at Smarts HSE (Health, Safety, Environment) Laboratory, University of Ulsan (UOU), have been more than fellow-travelers. Other than the apt and much-required critique and feedback of my work they have come with possible solutions and we have brainstormed together to solve different research problems. I would like to sincerely remember the help and collaboration in the early phase of research with my colleague Dr. Myeongsu Kang, an author of a bestselling book— Prognostics and Health Management of Electronics, who unexpectedly passed away at the age of 38. I am thankful to, no particular order, Jaeyoung Kim, Inkyu Jeong, Dong-gu Son, Nguyen Ngoc Hung, Prosvirin Alexander, Farzin Piltan and other lab mates for their valuable feedback and recommendation throughout the research seminars and always available when asked for help.

I would also like to gratefully acknowledge the financial support of the BK21 (Brain Korea 21) plus scholarship program, the Korea Institute of Energy Technology Evaluation and Planning (KETEP), the Ministry of Trade, Industry and Energy (MOTIE) of the Republic of Korea, the

National Research Foundation of Korea (NRF) and the Ministry of Science, ICT and Future Planning which allow me to pursue my study and research in South Korea.

M M Manjurul Islam

University of Ulsan, Ulsan, Republic of Korea

June 2019

ABSTRACT

Reliable fault diagnosis and prognosis (FDP) of complex engineering systems is a pressing need to prevent catastrophic failure by avoiding unanticipated problems that could lead to performance deficiencies and adverse effects on safety. In the era of the Internet of Things (IoT), the dramatic increase of sensors, data rates, and communication capabilities continue to drive the complexity of FDP applications to new levels. As a result, governments and commercial industrial communities are looking for new insights to use the massive volume of streaming in from their systems and sensors. Therefore, this dissertation presents a data-driven FDP framework for rotating bearings in large-scale industries based on advanced signal processing and improved machine learning (ML) techniques. This framework comprises of four important modules: robust condition monitoring scheme-based on time-frequency signal analysis, feature engineering (i.e. classical machine learning)- and feature learning (i.e. deep learning)-based reliable fault diagnosis methodology, and data-driven prognostics framework using new health index (HI) and variants of least-square support vector machines for remaining useful life (RUL) estimation.

To determine bearing health in operation for detecting an impending failure at an early stage, this dissertation proposes a robust condition mentoring methodology for bearing failures that employs time-frequency analysis (TFA) and optimum sub-band analysis on the stream of acoustic emission (AE) signals to select informative sub-bands of the signal. There is no general consensus on how many samples and which portion of the time-domain signal should be analyzed. To address this issue, wavelet packet transform-based envelope analysis with degree-of-defectiveness ratio (WPT-EA+ DDR) evaluation matrices is applied to quantify each sub-band signal. The results of (WPT-EA+ DDR) are visualized in two-dimensional (2D) analysis tool as a percentage of degree-of-defectiveness ratio (DDR) values. This 2D tool is highly effective to select a narrow-band signal from the stream of signal that contains the most intrinsic and pertinent information about the defects. To verify the effectiveness of the proposed (WPT-EA+ DDR) sub-band selection method, this method is compared with recent state-of-the-art methods.

To realize highly reliable data-driven fault diagnosis of a complicated industrial process for identifying the root cause of failures, it is important to exploit useful and discriminatory features from measured data. As a result, this dissertation proposes a hybrid feature selection (HFS) scheme for identifying the most discriminant fault signatures using an improved class separability criterion—the local compactness and global separability (LCGS)—of distribution in feature

dimension to diagnose bearing faults. The HFS model consists of filter-based selection and wrapper-based selection. In the filter phase, a sequential forward floating selection (SFFS) algorithm is employed to yield a series of suboptimal feature subset candidates using LCGS based feature subset evaluation metric. In the wrapper phase, the most discriminant feature subset is then selected from suboptimal feature subsets based on maximum frequency of occurrence and maximum average classification accuracy (ACA) estimation of support vector machine (SVM) classifier using them. The effectiveness of the proposed HFS method is verified with fault diagnosis application for low speed rolling element bearings (REBs) under various conditions. Experimental results indicate that the proposed method outperforms the state-of-the-art algorithm when selecting the most discriminate fault feature subset, yielding an improvement of diagnostics performances in average classification accuracy.

Furthermore, fault diagnosis in variable operating conditions is still a challenging problem since fault characteristics significantly vary with these changing conditions. An advent of deep learning (DL) has been introduced for fault diagnosis in most recent years. DL can automatically learn useful features from raw signals even with locally distorted and translated of characteristics of information. Therefore, this dissertation proposes a reliable fault diagnosis scheme under variable speed conditions based on adaptive deep convolutional neural networks (ADCNN) with acoustic spectrum imaging (ASI) of acoustic emission (AE) signals as a precise health state. In ASI, the amplitudes of the spectral components of the windowed time-domain acoustic emission signal are transformed into spectrum imaging. ASI provides a visual representation of acoustic emission spectral features in images. This ensures enhanced spectrum images for ADCNN testing and training, and thus provides a robust classifier technique with high diagnostic accuracy. To verify the proposed (ASI + ADCNN), benchmark bearing dataset with variable conditions are used.

Finally, this dissertation presents a data-driven prognostic framework for rolling-element bearings (REBs). This framework infers a bearing's health index by defining a degree-of-defectiveness (DD) metric in the frequency domain of bearing raw signal, named DD-based health index (DD-HI). Then, this dissertation systematically apply least-square support vector machines (LSSVMs) in the forms of Bayesian inference-aided one-class LSSVM (Bayesian-OCLSSVM) for anomaly detection in order to define the time to start (TTS) point of remaining useful life (RUL) prediction and the recurrent least-square support vector regression (Recurrent-LSSVR) model for predicting future values of DD-HI for calculating the RUL. In addition, this dissertation addresses several pertinent challenges, such as failure threshold determination during anomaly detection and RUL estimation, by developing adaptive thresholds. This research conducts extensive experiments

on a benchmark dataset using a run-to-failure experiment. The results demonstrate the efficacy of the proposed framework compared to state-of-the-art methods in terms of the accuracy and convergence of the RUL estimation of bearings.

Contents

VITA	viii
DEDICATION	ix
ACKNOWLEDGMENTS	x
ABSTRACT	xiii
Contents	xvii
List of Figures	xx
Nomenclature	xxiv
Chapter 1: Introduction	1
1.1 Background Information	1
1.2 PHM Methodologies	3
1.2.1 Canaries	3
1.2.2 Model-based Approaches	5
1.2.3 Data-Driven Approaches	6
1.3 Motivation and Objective of Research	7
1.3.1 Summary of Contributions	12
1.4 Dissertation Outlines	13
Chapter 2: Robust Condition Monitoring of Bearing Using Time-Frequency Signal Analysis and Optimum Sub-Band Selection	16
2.1 Introduction	16
2.2 Proposed Condition Monitoring Scheme for Low-Speed Bearings	18
2.2.1 Experiment Setup and AE data acquisition	19
2.2.2 WPT-EA with DDR to Select Informative Sub-Bands Regarding Bearing Defects	21
2.3 Results and Discussions	26
2.4 Conclusion.....	30

Chapter 3: Reliable Multi-Fault Diagnosis Scheme based on Heterogenous Feature Model and Hybrid Feature Analysis	31
3.1 Introduction	31
3.2 Experiment Setup and Acoustic Emission (AE) Signal Acquisition.....	35
3.3 Proposed Fault Diagnosis Methodology.....	37
3.3.1 Heterogeneous Feature Extraction Models (HFEM).....	37
3.3.2 Proposed HFS-LCGS Scheme.....	40
1) A feature evaluation metric to assess the quality of feature subset in HFS-LCGS	40
2) Accuracy estimation of the SVM classifier in the wrapper method of HFS-LCGS.....	42
3.3.3 Fault Classification for Online Diagnosis	43
3.4 Experiment Results and Discussion	44
3.5 Conclusions	50
Chapter 4: Acoustic Spectrum Imaging and Deep Convolution Neural Network for Rotating Bearings Fault Diagnosis Under Variable Operating Conditions.....	53
4.1 Introduction	53
4.2 Methodology	56
4.2.1 Experimental Testbed and Data Acquisition.....	57
4.2.2 Acoustic Spectrum Imaging (ASI) for Bearing Health State Visualization.....	58
4.2.3 Proposed Fault Diagnosis System with an ADCNN	61
1) Deep CNN (DCNN) Architecture Overview.....	61
2) An ADCNN with Back-propagation	61
4.3 Results and Discussion.....	65
4.3.1 Dataset description	65
4.3.2 Performance Analysis of Acoustic Spectrum Imaging (ASI).....	67
4.3.3 Diagnostic performance of the proposed method.....	67
4.4 Conclusions	72

Chapter 5: Data-Driven Prognostic Scheme for Rolling-Element Bearings Using a New Health Index and Variants of Least-Square Support Vector Machines	73
5.1 Introduction	73
5.2 Proposed Methodology.....	76
5.2.1 Degree-of-Defectiveness-based HI (DD-HI).....	77
5.2.2 Bayesian-OCLSSVM for Anomaly Detection	80
5.2.3 Recurrent Least-Square Support Vector Regression (Recurrent-LSSVR) for Predicting Future DD-HI Values and Estimating RUL	84
5.2.4 Failure Threshold.....	86
5.3 Bearing test rig and Run-to-failure dataset.....	87
5.4 Results verification and discussion	88
5.5 Conclusion.....	94
Chapter 6: Summary of Contributions and Future Work	96
6.1 Summary of Contributions	96
6.2 Future Work	98
Publications	101
International Journals	101
Under review	102
Book Chapters	102
International Conferences.....	103
References	105

List of Figures

Figure 1.1: A conceptual framework of PHM solution [1]	2
Figure 1.2: The “old” canary paradigm	3
Figure 1.3: (a) Canary fuse with cracked glass housing, (b) A filament-based canary precursor to predict failure [1].....	4
Figure 1.4: PoF-based prognostics approach.....	5
Figure 1.5: Data-driven fault diagnosis and prognosis approach.....	7
Figure 1.6: Cylindrical single row roller bearing with four essential components.	8
Figure 1.7: Faulty bearing signal for inner fault, outer fault and roller fault [16].....	9
Figure 1.8: Geometry of a rolling element bearing; where P_d is the pitch diameter, R_d is the roller diameter, and θ is the contact angle [27]	9
Figure 1.9: A typical Data-Driven FDP framework.....	10
Figure 2.1: A block diagram of the proposed condition monitoring scheme with sub-band analysis.	18
Figure 2.2: A screenshot of the self-designed fault simulator with a three-phase induction motor, a wide-band acoustic sensor, and a gearbox for fault diagnosis.....	19
Figure 2.3: Different bearing faults: a) BCO, (b) BCI, (c) BCR.....	20
Figure 2.4: Flowchart of the WPT-based envelope analysis (WPT-EA) with DDR to select informative sub-bands regarding bearing defects.....	20
Figure 2.5: Detailed framework of the DDR calculation for WPT-EA at each node.....	21
Figure 2.6: Two-dimensional visualization tool of the percentage of each $DDR_{overall}$ for selecting the wavelet filter	25
Figure 2.7: An example of 2D visualization of DDR values for BFO signal.....	26
Figure 2.8: Original AE signals of each bearing at speed conditions (a) BCO at 300 RPM, (b) BCI at 350 RPM and (c) BCR at 400 RPM	27
Figure 2.9: The results of FFT and its limitation for finding the defect peak. (a) An original AE signal of the BCO fault, (b) FFT of the signal in (a), and (c) enhanced view of (b) where there is no BPFO or its harmonics.	27
Figure 2.10: 2D visualization tool for finding informative sub-band signals based on (a) SKV and (b) the proposed DDR values	28
Figure 2.11: Original AE signals of different faults such as BCO, BCI, and BCR at 300 RPM (Left) and at 400 RPM (Right)	29

Figure 2.12: Effectiveness of the proposed method for detecting different bearing defects for (a) 300 RPM (left) and (b) 400 RPM (Right)	29
Figure 3.1: Screenshot of the self-designed experiment setup, (a) standard equipment setup, (b) PCI based AE system for data acquisition.	34
Figure 3.2: Different single and combined bearing faults with 3 mm crack: a) BCO, b) BCI, c) BCR, d) BCFIO, e) BCOR, f) BCIR, and g) BCIOR.....	35
Figure 3.3: Different single and combined bearing faults with 12 mm crack: a) BCO, b) BCI, c) BCR, d) BCFIO, e) BCOR, f) BCIR, and g) BCIOR.....	35
Figure 3.4: Acquired original AE signals of different bearing conditions in (a) dataset 1 and (b) dataset 2.....	36
Figure 3.5: An overall block diagram of improved bearing fault diagnosis scheme with HFS scheme.....	37
Figure 3.6: The decomposition method of WPT up to three levels for extracting energy features	39
Figure 3.7: Overall procedure of the developed HFS-LCGS scheme for selecting most discriminant feature subset.....	41
Figure 3.8: The conception of proposed evaluation metric—local compactness and global separability (LCGS) calculation, (a) local compactness (e.g. in 1D feature space) and (b) global separability, where black circles define the local compactness (e.g. in 2D feature space)	41
Figure 3.9: Estimated classification accuracy of SVM classifier for 30 feature subsets candidates	45
Figure 3.10: Estimated classification accuracy of SVM classifier for 30 feature subsets candidates	45
Figure 3.11: Estimated classification accuracy of SVM classifier for 30 feature subsets candidates	46
Figure 3.12: Estimated classification accuracy of SVM classifier for 30 feature subsets candidates	46
Figure 3.13: The results of confusion metrics of the proposed method to demonstrate actual vs. predicted deviation for four different datasets.....	48
Figure 3.14: Sensitivity and precession results of the proposed method for different fault classes	49
Figure 3.15: 2D visualization of most discriminant feature subset results by developed HFS-LCGS and the state-of-the-art for (a) dataset 1 and (b) dataset 4	50

Figure 4.1: The overall structure of the ADCNN-based method for RPM variant fault diagnosis.	56
Figure 4.2: (a) Experimental testbed, (b) PCI 2-based data acquisition system, and (c) Examples of ORC, IRC, and RRC bearing defects.	57
Figure 4.3: Flowchart of the acoustic spectrum imaging (ASI) for visualization of the bearing health condition.	59
Figure 4.4: Adjustable sliding window technique for the proposed acoustic spectral imaging (ASI).	60
Figure 4.5: The proposed ADCNN architecture used to classify different fault types, where an ASI containing the bearing health state feeds the deep neural network as the input data.	61
Figure 4.6: Different raw signals for different health conditions: (a) NC, (b) IRC, (c) ORC, and (d) RRC.....	66
Figure 4.7: Different ASI for different health conditions: (a) NC, (b) IRC, (c) ORC, and (d) RRC	66
Figure 4.8: 3-dimensional visualization of different features using KPCA: (a) [108], (b) [107], (c) [102], and (d) Proposed.	71
Figure 5.1: Proposed prognosis framework for estimating the remaining useful life (RUL).....	76
Figure 5.2: Bearing configuration with possible fault locations.	77
Figure 5.3: (a) An envelope power spectrum of a combined fault, and an illustration of the defect range calculation up to three harmonics (using equations (8), (9), and (10)) in the envelope power spectrum on which RMS features are calculated for (b) outer, (c) inner, and (d) roller defects, respectively.....	78
Figure 5.4: One-class, least-square support vector machine (OCLSSVM).....	80
Figure 5.5: Proposed recurrent training framework for least-square support vector regression. ..	86
Figure 5.6: Updating process of the proposed Recurrent-LSSVR model.	87
Figure 5.7: Experimental setup and component location for recording run-to-test failure data for bearing RUL estimation [29].....	88
Figure 5.8: Different health indexes of bearing run-to-test failure data for (a) bearing no. 2 and (b) bearing no. 4 using RMS, Kurtosis, and DD-HI (proposed).	90
Figure 5.9: Anomaly detector output for (a) bearing no. 2 and (b) bearing no. 4. The first 50 instances of the initial data are used for pre-training the anomaly detector.	90
Figure 5.10: Gradient-based dynamic thresholding method at various points. In this figure, blue indicates the true trend, green indicates the predicted trend, and red indicates the gradient at different points.	90

Figure 5.11: One-step-ahead prediction results of Recurrent-LSSVR for (a) bearing no. 2 and (b) bearing no. 4. After the TTS point, 50 instances are used for training to build the baseline model (red color)..... 91

Figure 5.12: RUL prediction performance of the proposed scheme. 92

Nomenclature

1D	One Dimensional
2D	Two Dimensional
ACA	Average classification accuracy
AE	Acoustic emission
ASI	Acoustic spectrum imaging
ANN	Artificial neural network
ASI+ADCNN	Acoustic spectrum imaging with Adaptive deep convolutional neural network architecture
Bayesian-OCLSSVM	Bayesian inference-aided one-class LSSVM
BCI	Bearing crack on inner raceway
BCIO	Bearing crack on inner and outer raceways
BFIOR	Bearing fault on inner raceway, outer raceway, and roller
BCIR	Bearing crack on inner raceway and roller
BCO	Bearing crack on outer raceway
BCOR	Bearing crack on outer raceway and roller
BCR	Bearing crack on roller element
BND	bearing with no defect
BP	Back-propagation
BSF	Ball spin frequency
CBM	Condition-based maintenance
CF	Crest factor
DBF	Deep belief networks
DCNN	Deep convolutional neural networks
DD-HI	Degree-of-defectiveness based health index
DL	Deep learning
DDR	Degree-of-defectiveness ratio
DNN	Deep neural network
FDD	Fault detection and diagnosis
FDP	Fault diagnosis and prognosis
FFB	Fault free bearing
FFT	Fast Fourier transformation
FMMEA	Failure Modes, Mechanisms, and Effects Analysis
GMM_{window}	Gaussian mixture model-based windows
GS	Global separability
HFEM	Heterogeneous feature extraction model
HFS	Hybrid feature selection
HI	Health index
IF	Impulse factor
IMS	Impulse factor
IoT	Internet of Things

<i>k</i> -cv	<i>k</i> -cross validation
<i>k</i> -NN	<i>k</i> -nearest neighbor
LC	Local compactness
LCGS	Local compactness and global separability
LSSVM	Least-square support vector machines
ML	Machine learning
NDES,	Non drive-end shaft
NM	Nearest mean
NN	Nearest neighbor
OAA	One-against-all
OAA-MCSVM	One-against-all Multi-class Support vector machine
OBJ	Objective function
PHM	Prognostics and health management
PoF	Physics of failure
PP	Peak-to-peak
REB	Rolling-element bearings
RMS	Root mean square
RPM	Revolution per minutes
RUL	Remaining useful life
RWE	Relative wavelet energy
SD	Standard Deviation
SF	Shape factor
SFFS	Sequential forward floating search
SGD	Stochastic gradient descent
SKV	Spectral kurtosis value
SOM	Self-organizing map
SNR	Signal to noise ratio
STFT	Short-time Fourier transformation
SVM	Support vector machine
TFA	Time-frequency analysis
WECSs	Wind energy conversion systems
WPT-EA	Wavelet packet transform-based envelope analysis
WT	Wind turbine

Chapter 1

Introduction

1.1 Background Information

With the rapid development of technology and demand for fast industrialization, rotating machinery has become more precise, larger and much more complicated. As a result, many catastrophic failures can occur in the rotating machinery, leading to unexpected breakdowns in the industry systems with consequences in tremendous losses in production, economic, and even human casualties. For instance, Toyota Motor North America recalled 28 600 model year 2018 and 39 900 model year 2012–2015 Prius Plug-In Hybrids due to parking brake malfunctioning in 2017 [1]. An aircraft engine caught fire as it accelerated for take-off as Korean Air Flight 2708 at Japan's Haneda Airport in May 2016. On July 23, 2011, two high-speed trains collided on a crossing in the suburbs of Wenzhou, Zhejiang province, China, resulting in human casualties [1]. According to the official investigation, the accident was caused by faulty signal systems that failed to warn the second train of the stationary first train on the same track [1].

All these incidents could have been prevented if there was an intelligent fault diagnostic prognostics (FDP) system. The main motivation for specifying more advanced diagnostic and prognostic requirements is the realization that they are needed to fully enable and obtain the benefits of new and revolutionary logistic support concepts and technology. These logistic support concepts are called by many names and include fault diagnosis and prognosis (FDP), condition-based maintenance (CBM), and autonomic logistics—all of which comprise prognostics and health management (PHM) capabilities as a key enabler [2]. PHM is a multilayered discipline that shields the integrity of components, products, and systems of systems by avoiding unexpected problems

that can lead to performance deficiencies and adverse effects on safety. More specifically, prognostics is the process of estimating a system's remaining useful life (RUL) by predicting the future behavior of a fault given the current degree of degradation and the load history.

PHM involves sensors and sensing strategy, anomaly detection via in-situ condition monitoring, diagnostics, prognostics, and decision support, as depicted in Figure 1.1. Sensing strategy is to collect data of time-dependent operation of a system, the degradation of materials, and/or the environmental loads on the components of a system or the entire product.

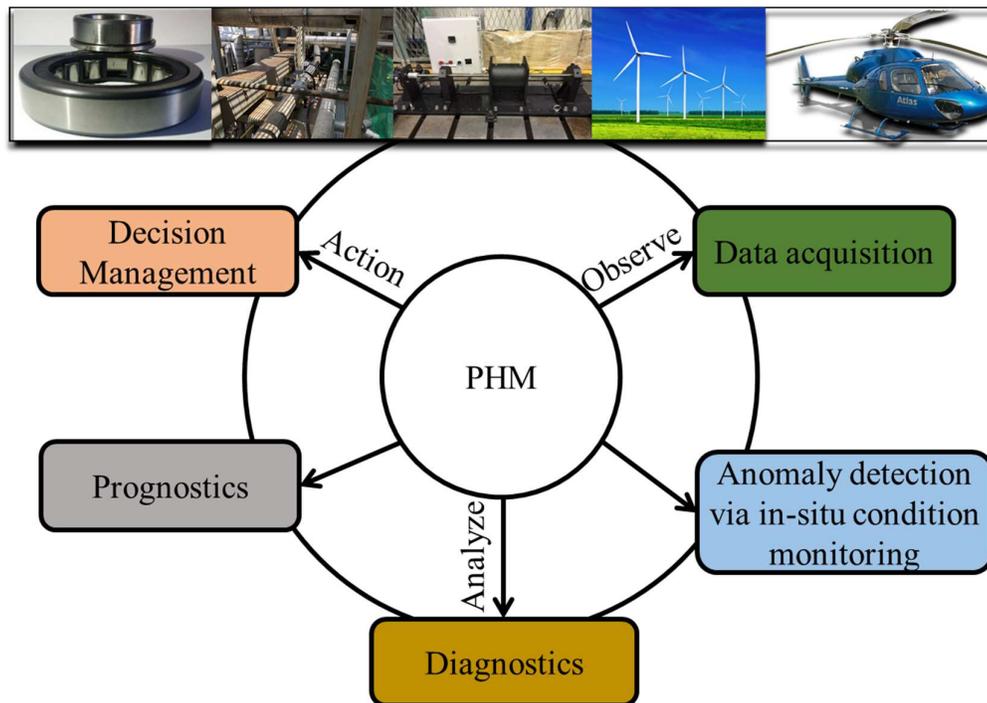


Figure 1.1: A conceptual framework of PHM solution [1]

The main purpose of anomaly detection is to detect unusual (e.g., anomalous) behavior of the system by identifying deviations from nominally healthy behavior. The results from anomaly detection can provide advanced warnings of an impending failure, often referred to as failure precursors. Note that anomalies do not necessarily indicate failure because changes in operating and environmental conditions can influence sensor data to show anomalous behavior. However, even this type of anomaly information is valuable to product health management, since it can indicate an unexpected use.

Diagnostics module enables the extraction of specific fault-related information, such as failure types, fault locations, the severity of the damage, and so forth, from sensor data caused by anomalies in the health of the product. This is a key piece of information that feeds into maintenance planning and logistics. Prognostics specifies the process of estimating a system's RUL within certain confidence intervals, which often requires additional information not traditionally provided by sensors, such as maintenance history, past and future operating profiles, and environmental factors. Based on prediction results, the goal is to inform decision-makers of potential cost avoidance activities and to ensure safe operation. Thus, the aspects of FDP/PHM are to effect appropriate decision-making; to prevent catastrophic system failures; to increase system availability by reducing downtime; to expend maintenance cycles; to execute timely repair actions; to lower life-cycle costs by reductions in inspection and repair; and to improve system qualification, design, and logistical support.

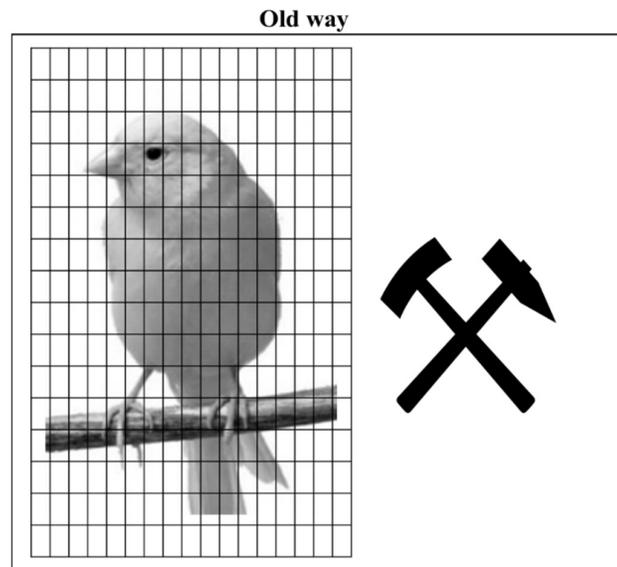


Figure 1.2: The “old” canary paradigm

1.2 PHM Methodologies

In compare with past times, PHM capabilities are evolved and matured. Canary, model-based and data-driven approaches have been studied to allow reliable and robust FDP.

1.2.1 Canaries

In the “old” days, a canary went down to work with coal miners and was monitored periodically for maintenance and decision support, as shown in Figure 1.2. If the canary died, they

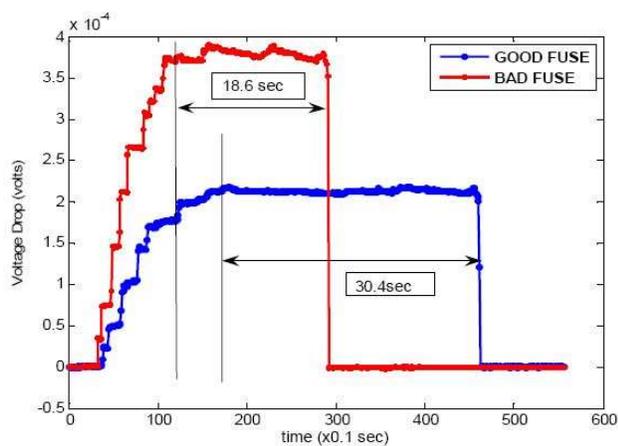
knew that they were soon to follow if they didn't run [2]. That is, due to the ability that canaries can detect deadly gas in the coal mine, they were used to be indicators giving an alarm.

An emulation of canary using a filament can be explained as follows and depicted in Figure 1.3 for industries.

- Filaments are used in light bulbs and in X-Ray tubes to utilize electrical energy to generate photons of the desired frequency range.
- However, some of the input electrical energy is also converted to thermal energy which results in long-term degradation and failure of the filament.
- Hence, an error-seeded canary that is designed to age faster than the functional filament is used as a precursor to predicting failure.



(a)



(b)

Figure 1.3: (a) Canary fuse with cracked glass housing, (b) A filament-based canary precursor to predict failure [1]

1.2.2 Model-based Approaches

The term ‘model-based’ or PoF (physics of failure) is interchangeable in literature. In general, the PoF approach involves Failure Modes, Mechanisms, and Effects Analysis (FMMEA) for life-cycle load monitoring, load feature extraction, and damage assessment. FMMEA is a systematic methodology to identify potential failure mechanisms and models for all potential failure modes and to prioritize failure mechanisms, as shown in Figure 1.4.

- **Life-Cycle Load Monitoring**
 - The life-cycle environment consists of manufacturing, shipment, storage, handling, and operating and non-operating conditions, which leads to performance or physical degradation of the system.
 - Hence, *in-situ* life-cycle load monitoring is needed to design practical damage models, which account for degradation due to cumulative load exposures.

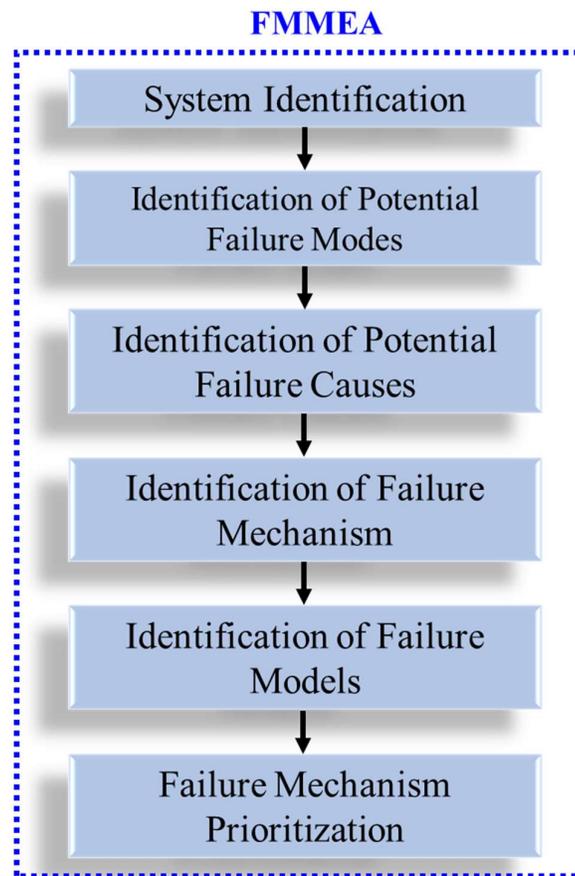


Figure 1.4: PoF-based prognostics approach.

- **Load Feature Extraction**
 - To reduce a large volume of load parameters but to maintain useful parameters for further use in assessing performance degradation.
- **Damage Assessment and Physics-Based Models Used for RUL Forecasting**
 - Temperature and vibration are common load conditions that can accelerate failure prediction.
 - Assess temperature- and vibration-induced damage.

1.2.3 Data-Driven Approaches

In data-driven techniques, data analytics and machine learning are used to determine anomalies and make predictions about the reliability of systems. A systematic data-driven approach for fault diagnosis and prognosis is given in Figure 1.5.

- **Diagnosis and Prognosis**
 - Dimensionality reduction (DR)/Feature Selection (FS): Multivariate statistical methods [3, 4], static and dynamic principal component analysis [5, 6], linear and quadratic discriminate, canonical variated analysis[7] are widely used.
 - Clustering: k-means, fuzzy c-means, self-organizing map [8]
 - Classification: black box methods based on neural networks [9], decision trees, support vector machines [10, 11]
 - Prediction: hidden Markov model, Kalman filter [12, 13], particle filter
- **Advantages**
 - Compared to PoF approaches, data-driven approaches do not necessarily need system-specific information.
 - Analyze intermittent faults by detecting changes in system features.
 - Can be adopted in complex systems.

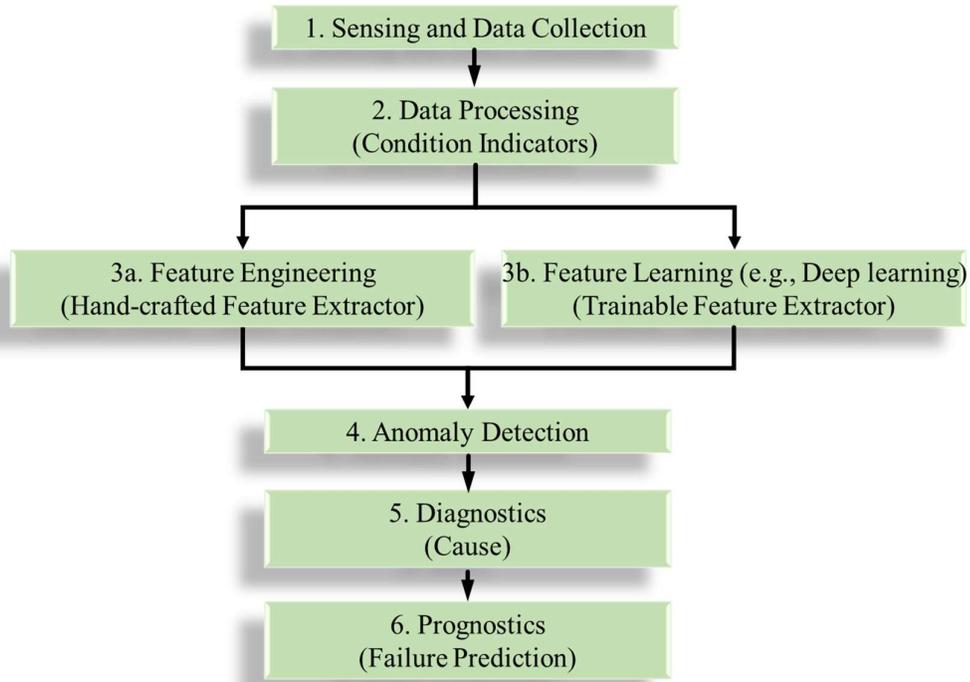


Figure 1.5: Data-driven fault diagnosis and prognosis approach.

1.3 Motivation and Objective of Research

Rotating bearings are considered as the most critical components for carrying heavy loads and providing constant rotational speed in the industrial machines. These bearings can be found in various industrial components from induction motors [14-17], gearboxes [18, 19], locomotives, aircraft, automotive, turbine engines [17, 20], spacecraft [21, 22]. The failures of rotating machinery are generally classified into four main groups: bearing faults, rotor-related faults, stator-related faults, and other faults (cooling, connection, terminal boxes) [3, 16]. According to other study conducted by Electric Power Research Institute (EPRI) and the Motor Reliability Working Group of IEEE-IAS, more than 50% of critical failures in rotating machinery are related to the rolling element bearings [23]. Bearing faults are also regarded as the main cause of equipment malfunctions and unplanned downtimes of rotating machinery in electrical drives and industrial plants [24, 25]. A rolling element bearing (REB) typically consists of two concentric raceways, outer raceway, and inner raceway with a set of rolling elements (or balls), which run in their tracks, as illustrated in Figure 1.6. Typically, the rolling elements are bound in a cage in order to maintain a constant angular pitch between the adjacent rolling elements and avoid unexpected mutual contact.

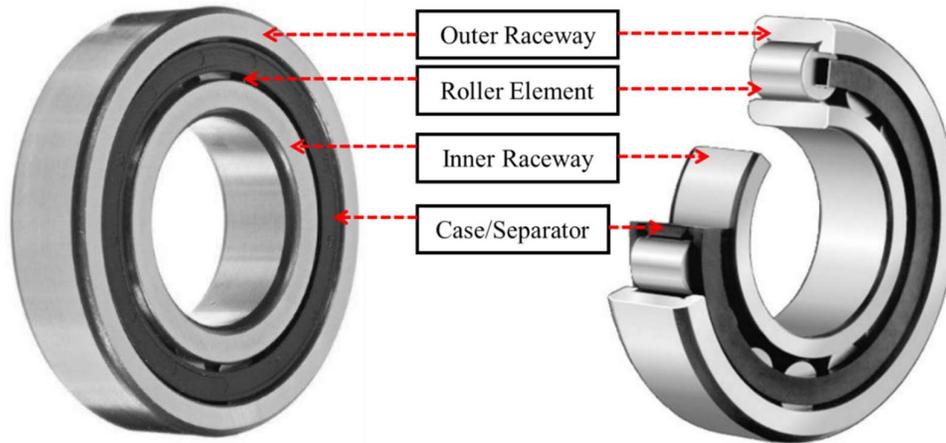


Figure 1.6: Cylindrical single row roller bearing with four essential components.

To explain bearing dynamics, a fault-free bearing (FFB) can also generate vibration or emit energy while it is rotating at a constant rotational speed because of a load, many static forces (e.g., preloading, weight of shafts, other moving components), and dynamic forces (i.e., centripetal forces, frictional forces, traction forces, fluid pressure, and some others) [26]. Nevertheless, when the bearing elements experienced some defects on the surface, some abnormal vibration can be observed due to contact with other components of bearing, which can create some abnormal oscillatory components in the signal. According to bearing dynamics, these oscillatory waves are highly nonlinear and nonstationary signals and depend on localizing defects (e.g., outer raceway, inner raceway, roller, and case faults). Figure 1.7 depicts various defect signals while rolling strikes a localized defect.

In Figure 1.7, there are three faults such as outer raceway (red dot), inner raceway (blue dot), and roller element (green dot). To explain these faults signals, when the rolling element passes over the defect on the outer raceway, it generates a pulse in the waveform. The magnitudes of these impulses are the same since the outer raceway is stationary according to the load zone area and the position of the sensor. Similarly, when the rolling element passes over the defect on inner raceways, it generates a pulse on the waveform. However, the inner raceway is rotating, and the magnitudes of impulses are varying according to the position of fault at load and no-load zone. Similarly, the fault on roller hits to inner and outer race and generates a pulse and the magnitudes of impulses are varies according

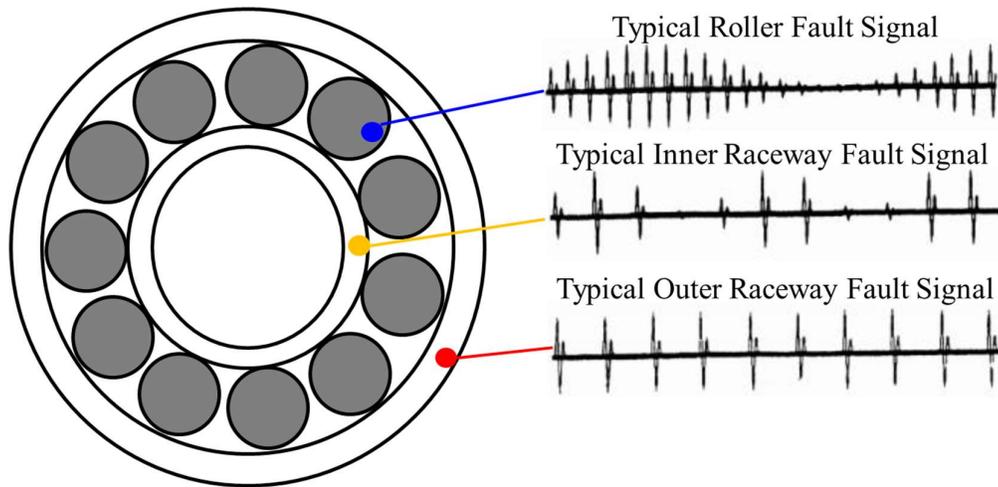


Figure 1.7: Faulty bearing signal for inner fault, outer fault and roller fault [16].

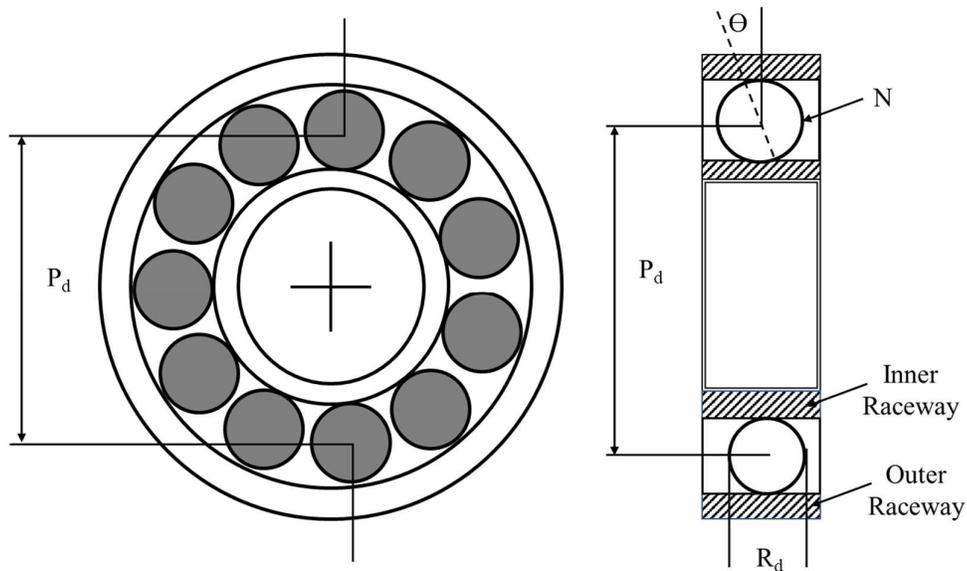


Figure 1.8: Geometry of a rolling element bearing; where P_d is the pitch diameter, R_d is the roller diameter, and θ is the contact angle [27]

to load and non-load zone [16]. The resulting impulse of vibration or emitted energy repeats periodically at a rate, which called the bearing fault frequency. The fault characteristic frequencies can be revealed by the location of the fault and the geometry of bearing and defined in Eq. (1.1). Figure 1.8 represents the geometry of a rolling element bearing. Four distinguishable fault frequencies are observed and those depend on the number of rollers, N , the shaft speed, S_{shaft} , in

hertz, the contact angle θ , the roller diameter, R_d , and the pitch diameter, P_d . The bearing defect frequencies, i.e., roller pass frequency on the outer raceway (RPFO), roller pass frequency on the inner raceway (RPFI), roller spin frequency (RSF), and fundamental train frequency (FTF), can be calculated as follows:

$$\begin{aligned}
 RPFO &= \frac{N \cdot S_{shaft}}{2} \left(1 - \frac{R_d}{P_d} \cos \theta \right), \\
 RPFI &= \frac{N \cdot S_{shaft}}{2} \left(1 + \frac{R_d}{P_d} \cos \theta \right), \\
 RSF &= \frac{P_d \cdot S_{shaft}}{2 \cdot R_d} \left(1 - \left(\frac{R_d}{P_d} \cos \theta \right)^2 \right), \text{ and} \\
 FTF &= \frac{S_{shaft}}{2} \left(1 - \frac{R_d}{P_d} \cos \theta \right),
 \end{aligned} \tag{1.1}$$

During the past several decades, model-based FDP techniques have been extensively used for enhancing the reliability and availability of rotating bearing the industrial system subject to faults. However, today's industrial processes have become more complicated with less tolerance for both performance deterioration and productivity decrease.

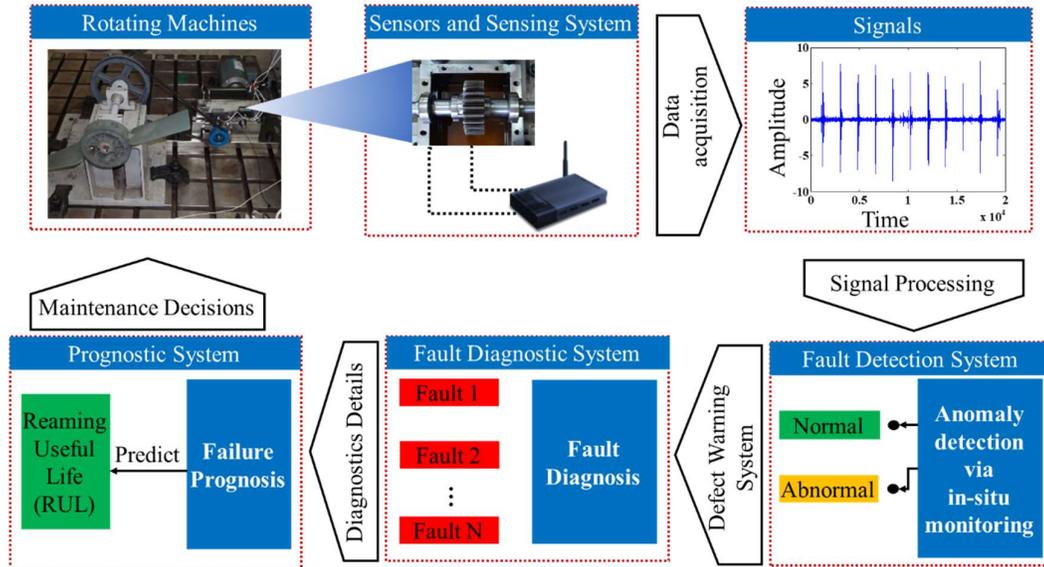


Figure 1.9: A typical Data-Driven FDP framework.

This implies that conventional model-based schemes demanding a deep knowledge of process model that can be derived from the first principles have been impractical for a complicated

industrial process. Fortunately, the rapid development of data mining, data acquisition techniques, machine learning, and high-performance techniques since 1990 has provided the ability collect and store a vast amount process data to extract useful information inherent in such a significant amount of recorded data (e.g., vibration, acoustic emission (AE), and current) for the purpose of reliable fault diagnosis and prognosis in the large-scale industries [4, 28].

Accordingly, this dissertation proposes a data-driven FDP framework for rotating bearing using advanced signal processing and machine learning techniques, as depicted in Figure 1.9. In general, data-driven FDP schemes involve sensors and sensing strategy for data acquisitions, fault feature calculation and decision-making procedure for diagnosis and prognosis and these are briefly summarized as follows,

- **Data Acquisition**

The primary goal of this process is to develop hardware technologies and sensing strategies for acquiring FDP relevant data that are promising to improve substantially the reliability, cost-effectiveness, coverage, and sensitivity of the fault signatures of the monitoring and fault tracking devices. These data could be the time-domain history of machines health condition such as acoustic emission (AE) and vibration.

- **Signal Processing for Feature Calculation**

Statistical parameters, which can be calculated from the measured time-domain data to represent conditions of the critical component in the industrial equipment, have been widely used a fault feature. The followings are representative statically parameters used for FDP: peak-to-peak (PP), standard deviation (SD), root-mean-square (RMS), kurtosis, skewness, crest factor (CF), impulse factor (IF), and shape factor (SF). In addition, time-frequency analysis (TFA) using short-time Fourier transform (STFT), wavelet transform, empirical mode decomposition, and cyclo stationary are frequently used for dealing nonstationary properties of bearing fault signals.

- **Machine learning (ML) for Decision-making**

Either supervised (e.g., support vector machine, artificial neural network and adaptive nearest neighbor, convolutional neural network) or unsupervised learning

classification methodology are widely employed to predict the type, size, location of defects, and RUL that can occur in the industrial system.

1.3.1 Summary of Contributions

The contributions of this dissertation are summarized as follows:

- A robust condition mentoring methodology for bearing failures using time-frequency signal analysis (TFA) and optimum sub-band analysis is proposed to select informative sub-bands of the signal. There is no general consensus on how many samples and which portion of the time-domain signal should be analyzed. To address this issue, wavelet packet transform-based envelope analysis with degree-of-defectiveness evaluation matrices (WPT-EA+ DDR) is applied to quantify each sub-band signal. The results of (WPT-EA+ DDR) are visualized in two-dimensional (2D) analysis tool as a percentage of degree-of-defectiveness values. This 2D tool is highly effective to select a narrow-band signal from the stream of signal that contains the most intrinsic and pertinent information about the defects.
- To realize highly reliable data-driven fault diagnosis of complicated industrial process for identifying root cause of failures, this dissertation proposes a hybrid feature selection (HFS) scheme for identifying the most discriminant fault signatures using an improved class separability criterion—the local compactness and global separability (LCGS)—of distribution in feature dimension to diagnose bearing faults. The HFS model consists of filter-based selection and wrapper-based selection. In the filter phase, a sequential forward floating selection (SFFS) algorithm is employed to yield a series of suboptimal feature subset candidates using LCGS based feature subset evaluation metric. In the wrapper phase, the most discriminant feature subset is then selected from suboptimal feature subsets based on the highest frequency of occurrences and maximum average classification accuracy estimation of support vector machine (SVM) classifier using them. The effectiveness of the proposed hybrid feature selection method is verified with fault diagnosis application for low speed rolling element bearings under various conditions.
- Furthermore, fault diagnosis in variable operating conditions is still a challenging problem since fault characteristics significantly vary with these changing

conditions. Therefore, this dissertation proposes a reliable fault diagnosis methodology under variable speed condition using deep convolutional neural networks (DCNN) with acoustic spectrum imaging (ASI) of acoustic emission (AE) signals as a precise health state. In ASI, the amplitudes of the spectral components of the windowed time-domain acoustic emission signal are transformed into spectrum imaging. ASI provides a visual representation of acoustic emission spectral features in images. This ensures enhanced spectrum images for DCNN testing and training, and thus provides a robust classifier technique with high diagnostic accuracy. To verify the proposed (ASI + ADCNN), benchmark bearing dataset with variable conditions are used.

- A data-driven prognostic framework for rolling-element bearings (REBs) is presented for robust remaining useful life (RUL) estimation. This framework infers a bearing's health index by defining a degree-of-defectiveness (DD) metric in the frequency domain of bearing raw signal, named DD-based health index (DD-HI). Then, least-square support vector machines (LSSVMs) is applied in the forms of Bayesian inference-aided one-class LSSVM (Bayesian-OCLSSVM) for anomaly detection in order to define the time to start (TTS) point of RUL prediction and the recurrent least-square support vector regression (Recurrent-LSSVR) model for predicting future values of DD-HI for calculating the RUL. In addition, this dissertation addresses several pertinent challenges, such as failure threshold determination during anomaly detection and RUL estimation, by developing adaptive thresholds. A set of experiments is conducted on a benchmark dataset of run-to-failure to verify the efficacy of the proposed framework.

1.4 Dissertation Outlines

The contents of this dissertation are divided into four main chapters, excluding Introduction and Summary and Future Work chapters. The contents of each chapter are briefly outlined below:

Chapter 1 introduces the necessary background information regarding condition monitoring, fault diagnosis, and fault prognosis through literature review, research objectives, and contributions of this dissertation. Moreover, it also provides a brief introduction to rolling element bearings and the type of localized defects in bearings.

Chapter 2 proposes a condition monitoring scheme for early identification of low-speed rolling element bearing defects using wavelet packet transform-based envelope analysis (WPT-EA) with degree-of-defectiveness ratio (DDR)-based sub-band analysis techniques. First, WPT-EA generates a set of sub-band signal and then DDR matrix is applied on each sub-band to quantify the degree of defectiveness. Finally, the performance of the proposed (WPT-EA + DDR) sub-band selected method is verified in a 2D visualization tool. This chapter is published in the *Journal of Ambient Intelligence and Humanized*.

Chapter 3 explores the feature engineering process that performs a diverse feature extraction process and hybrid feature selection (HFS) schemes for selecting the best feature subset for reliable fault diagnosis scheme. This diverse feature extraction process that tries to extract as many as features about fault that comprise of time-, frequency-, and time-frequency-domain signal analysis. These features are further refined applying HFS algorithm for selecting the best subset features. The proposed feature selection algorithm is varied using various single and multiple-combined faults of bearing fault diagnosis application. This chapter is published in the *Lecture Notes Series in Artificial Intelligence*.

Chapter 4 presents a feature learning process using the deep convolutional neural network (DCNN) for diagnosing various bearing faults under variable operating conditions. To utilize the advantage of the DCNN architecture, this chapter develops a novel signal-to-image technique for visualizing bearing fault information under variable operating speeds. These images are further utilized with DCNN network for fault diagnosis. This chapter is published in the *Measurement* journal.

Chapter 5 proposes a novel approach for the health prognosis of rolling element bearings using new health index (HI) and variants of least-square support vector machines for estimating RUL. To infer the HI of the bearing, this chapter explores the characteristic frequencies of the bearing defects and then quantities these frequencies using an evaluation matrix. Then, Bayesian-OCLSSVM is applied to detect anomalies in the constructed HI and Recurrent-LSSVR is used to estimate RUL. The performance of the proposed framework is validated using publicly available run-to-failure bearing dataset from Intelligent Maintenance System (IMS) center in the University of Cincinnati [29]. This chapter is under review in the *IEEE Transaction on Industrial Electronics*.

The dissertation is concluded in Chapter 6 with a summary of the contributions made therein and a discussion of future research directions.

Chapter 2

Robust Condition Monitoring of Bearing Using Time-Frequency Signal Analysis and Optimum Sub-Band Selection

2.1 Introduction

The energy crisis as well as environmental pollution, two major and pressing global environmental problems, are both driving the development of renewable energy sources. Wind energy conversion systems (WECSs) are becoming more popular as an economically viable alternative to fossil-fuel based power generation. WECS consisting of thousands of units are now forming a major portion of devices with renewable electrical generating capacity and play a vital role in the future of global energy generation [17, 30]. As the size of WECS continues to increase, their high maintenance costs and associated failure costs become increasingly important issues. A WECS usually includes four main components: the wind turbine (WT), generator, control systems, and an interconnection apparatus [17]. Among them, the WT is the element that fails most frequently [30]. More specifically, bearing defects account for the highest percentage of all failures in wind turbines because of their contribution to coarse operating conditions and other external influences, such as the ratio of high torque to low-speed, vibration, improper loading, and misalignment [3, 17]. Because abrupt mechanical failures affected by the bearing faults in WECSs result in a substantial economic loss, thus, reliable condition monitoring for detecting bearing failure at the early stages is crucial to prevent potential failures from turning into functional failures.

Motor current signature analysis [31], vibration techniques [32], temperature tests [31], and wear debris analysis have traditionally been used in the detection of bearing faults and have shown improved performance over time. A wide range of research has been carried out considering vibration, especially for diagnosing high-speed machinery [31, 33]. Acoustic emissions (AE)

detection is the latest technique in diagnosing faults in rolling element bearings (REBs) [34, 35]. The principal advantage of the AE technique over traditional vibration detection is that the former has a much better signal to noise ratio (SNR), even at very low frequencies, making it particularly suitable for detecting a possible failure at a very early stage. This study, therefore, records AE signals to detect faults in the early stages of crack development under the low-speed operation of rolling element bearings.

Whenever any defect such as a crack or spoil occurs on any of the four different bearing elements (*i.e.*, outer raceway, inner raceway, rollers or balls, and train or cage), it creates harmonics among the bearing characteristic (defect) frequencies (*e.g.*, Ball Pass Frequency Inner raceway (BPFI), Ball Pass Frequency Outer raceway (BPFO), and Ball Spin Frequency (BSF)) for each shaft rotation [32, 36]. It is important to note that a bearing defect symptom is hardly found around the raw and unfiltered harmonics of the defect frequencies in the original fault signal's power spectrum since it is an inherently nonstationary and nonlinear signal [15, 32]. Frequency analysis and demodulation are further needed. This overall process, called envelope analysis [32], focuses on the transient, impact-type events (spikes on the time domain signal) such as BPFI, BPFO and BSF, while the fast Fourier transform (FFT) process otherwise misses such transient events because of the way the FFT processes inherently nonstationary and nonlinear fault signals. Several researchers, therefore, explored impacts in the envelope power spectrums acquired from various sub-band signals using either short-time Fourier transform (STFT)[14, 37] or multi-level bandpass filters [15, 38].

Another important issue is to select pertinent and informative sub-band signals, one of the key contributions of this study, from a large input signal. The sub-band signals are further utilized for indemnifying of defects. In [38], Wang et al. recently introduced a wavelet-based kurtogram as a time-frequency analysis, which is broadly used to find useful sub-band signals since it can quantify the magnitudes of the rolling bearing defect frequencies BPFI, BPFO and BSF as well as their harmonics. However, this quantifying parameter is not precisely proportional to the degree of defectiveness of bearing rolling elements. To solve this problem, this chapter considers a Gaussian mixture model-based degree of defectiveness ratio (DDR) calculation, which is a ratio of defect-components to residual-components, instead of merely using a kurtosis value [38] in the envelope power spectrum of wavelet packet transform (WPT) nodes. The main concept of the DDR calculation is that it first generates a Gaussian window around the BPFI, BPFO, and BSF as well as their harmonics, and then calculates the DDRs about these defect frequencies. This evaluation metric provides a very efficient and meaningful way to accurately measure the degree of

defectiveness. Further, the highest DDR values about BPFI, BPFO, and BSF of the 2D visualizations of WPT nodes are selected as the most informative sub-bands.

2-D visualization based sub-band selections are apparently effective for finding appropriate fault conditions, which searches the inherently nonstationary and nonlinear AE signals via wavelet packet transform-based envelope analysis (WPT-EA) and selects the useful sub-band signals based on the highest DDR value in the 2-D visualization tool. In summary, the proposed robust condition monitoring scheme for bearings includes the following two essential steps: 1) condition monitoring by detecting the harmonics of the characteristic frequency in the WPT-EA. To quantify these characteristic frequencies in the WPT sub-band signal, DDR evaluation matrix is proposed. In addition, this chapter explores the impact of the Daubechies wavelet family during the calculation of DDR in order to evaluate the degree of defectiveness for various bearing failures, and 2) a 2D tool is used to visualize as a percentage of DDR values for identifying defects.

The remaining sections of this chapter are structured as follows. the proposed condition monitoring methodology including data acquisition system and WPT-EA with DDR-based sub-band analysis are presented in Section 2.2, results and discussions are given in Section 2.3, and finally concluding remarks are given in Section 2.4.

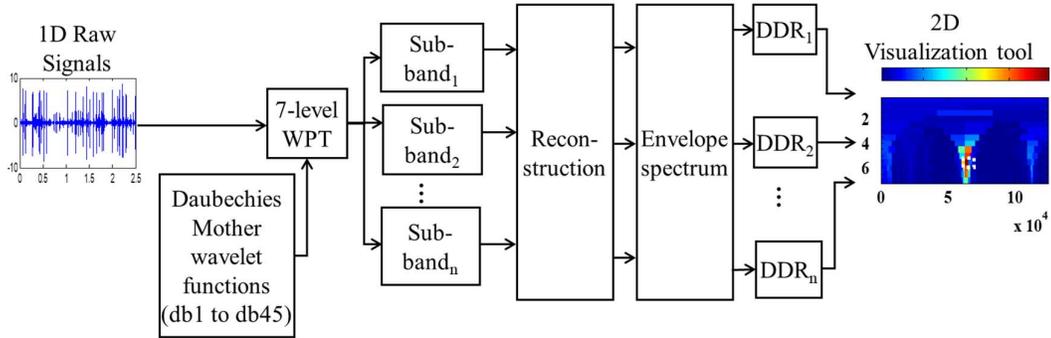


Figure 2.1: A block diagram of the proposed condition monitoring scheme with sub-band analysis.

2.2 Proposed Condition Monitoring Scheme for Low-Speed Bearings

Figure 2.1 illustrates an overall block diagram of the proposed condition monitoring methodology including WPT-EA and DDR-based sub-band selection.

2.2.1 Experiment Setup and AE data acquisition

A self-designed machinery fault simulator that has been used for proposed condition monitoring data acquisition, as shown in Figure 2.2. Acoustic emissions (AEs) is a popular method for detecting and monitoring low-speed bearings faults due to its powerful capability to capture dynamic activities in low-energy signals [34, 39-42]. AE signals were collected for bearings at the non-drive end of the simulator using a wide-band acoustic sensor (type WS α from Physical Acoustics Corporations [43]) and a PCI-2 based data acquisition system, which samples the AE signals at a rate of 250 KHz. The acoustic sensor is coupled to the top of the bearing housing, approximately 21.48 mm from the bearing, as can be seen in Figure 2.2. The non-drive end shaft is connected to the drive end through a gearbox with a reduction ratio of 1:1:1.52. The details of our experimental setup and data acquisition model can be studied in [11].

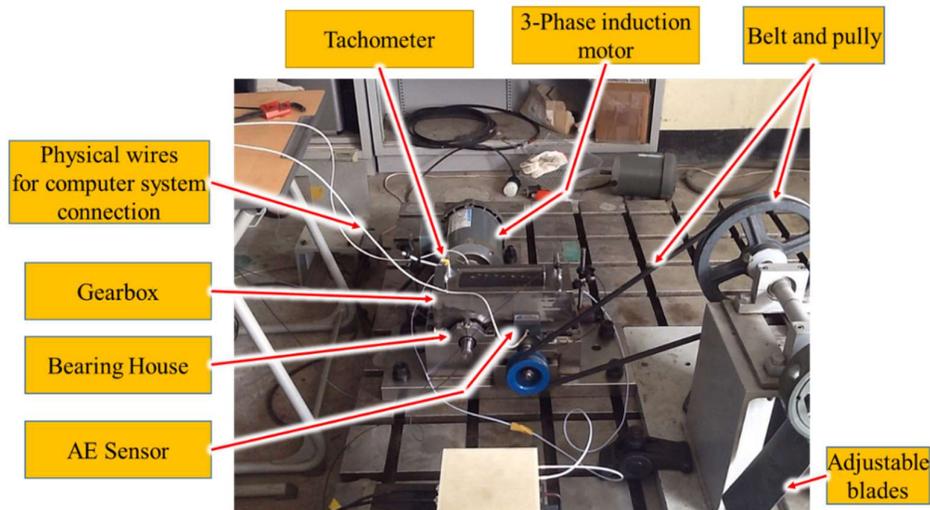


Figure 2.2: A screenshot of the self-designed fault simulator with a three-phase induction motor, a wide-band acoustic sensor, and a gearbox for fault diagnosis

In this chapter, cylindrical bearings (FAG NJ206-3-TVP2) were used for experiments, and all seeded single-point bearing defects include a bearing with a crack on its outer raceway (BCO), a bearing with a crack on its inner raceway (BCI), and a bearing with a crack on its roller (BCR). Figure 2.3 illustrates the single point seeded bearing defects. Table 2.1 presents the data descriptions that are used in this study.

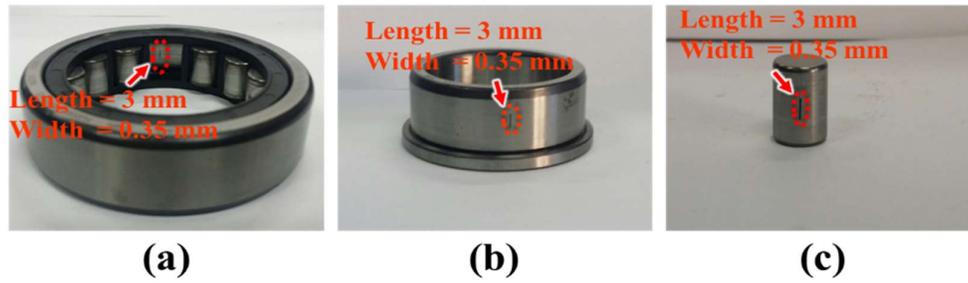


Figure 2.3: Different bearing faults: a) BCO, (b) BCI, (c) BCR

Table 2.2: Detailed description of seeded bearing defects including their characteristic frequencies

Seeded bearing defects	BCO	BCI	BCR
Revolutions-per-minute (RPM)	300	350	400
Length	3 mm	3 mm	3 mm
Width	0.35 mm	0.35 mm	0.35 mm

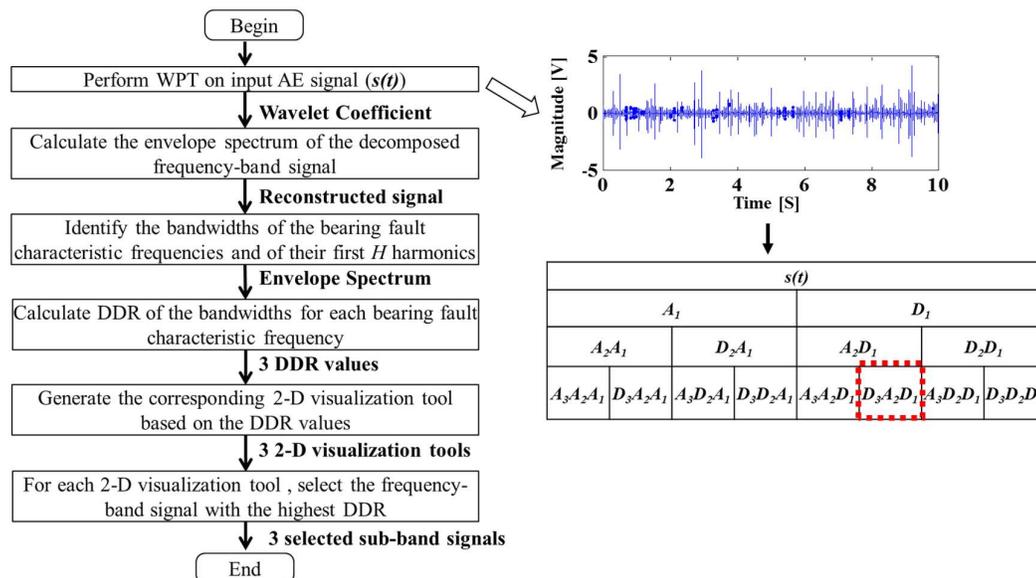


Figure 2.4: Flowchart of the WPT-based envelope analysis (WPT-EA) with DDR to select informative sub-bands regarding bearing defects

2.2.2 WPT-EA with DDR to Select Informative Sub-Bands Regarding Bearing Defects

As explained in Section 2.1, characteristic bearing defect frequencies are not observable in raw AE signals. Therefore, it is important to define an appropriate visualization tool that reveals a unique pattern about a bearing's health state. This chapter uses a WPT [44, 45] to explore the input signal and quantify each sub-band signal. WPT has been widely employed for analyzing the defect information inherent in the bearing fault signals due to its decomposition ability to split into low- and high-frequency sub-band signals. As illustrated in Figure 2.1, a bearing's AE signal is decomposed into a series of sub-bands by applying the WPT on n decomposition levels (i.e., $n=7$ in this study) using an appropriate mother wavelet function [43]. The 7-level WPT decomposition

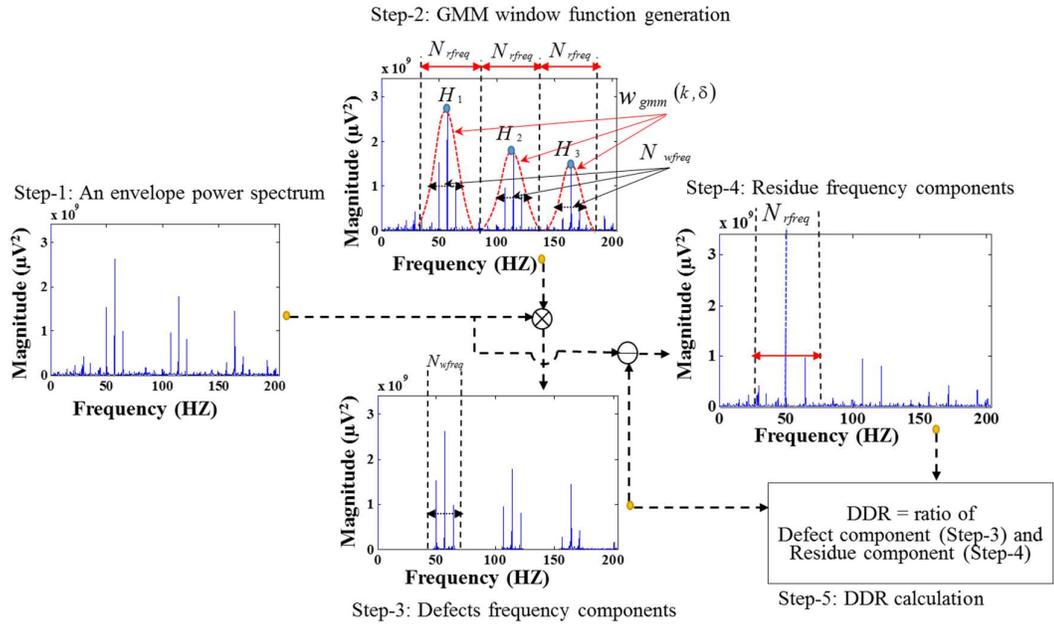


Figure 2.5: Detailed framework of the DDR calculation for WPT-EA at each node

generates 128 sub-band signals. Then, we calculate the DDR values for each node (e.g., $Node_{n,2^n}$ right-side of the figure) to quantify each informative sub-band, as can be seen in Figure 2.4. For accurate DDR calculation, we carefully analyze bearing defect (or characteristic) frequencies such as a ball pass frequency for the outer raceway (BPFO), a ball pass frequency for the inner raceway (BPFI), twice ball spin frequency ($2 \times BSF$), or a fundamental train frequency (FTF), which are observable for bearing failure. This calls for frequency analysis and a demodulation technique such as envelope analysis [11, 43] to reveal fault symptom. These DDR values precisely quantify the degree of defectiveness about the fault type (BPFI, BPFO, $2 \times BSF$). Thus, the 2D visualization of

those DDR values is used for condition-based monitoring. The detailed derivation of the DDR calculation is given below and illustrated in Figure 2.5.

Step-1. To compute an envelope power spectrum, the analytical signal of each segmented node of the WPT decomposition is calculated. For example, if $x(t)$ is the signal of the segmented node, the analytical signal can be represented as a combination of the original signal and the Hilbert transform [43]. The analytical signal can be formulated as below:

$$x^h(t) = x(t) + i\tilde{x}(t). \quad (2.1)$$

where $\tilde{x}(t)$ is the Hilbert transform. Convolution of the original signal with $1/\pi t$ yields the following:

$$\tilde{x}(t) = x(t) * \frac{1}{2\pi} = \frac{1}{\pi} \int_{-\infty}^{\infty} h(t) \frac{dt}{t - \tau}. \quad (2.2)$$

The power spectrum of the analytical signal is attained by calculating the square of the absolute value of the Fourier transform, defined as $\left| F\{x^h(t)\} \right|^2$, as can be seen in Step-1 of Figure 2.5.

This envelope spectrum in Eq. (2.2) is highly efficient at which faulty defect frequency must be observable, which can be defined as one of the following four defects (e.g., characteristics) frequencies depending on localizing defects (e.g., outer raceway, inner raceway, or roller raceway fault) [15, 43]:

$$\begin{aligned} BPF0 &= \frac{N.F_r}{2} \left(1 - \frac{R_d}{Pd} \cos \theta \right) \\ BPF1 &= \frac{N.F_r}{2} \left(1 + \frac{R_d}{Pd} \cos \theta \right) \\ BSF &= \frac{Pd.F_r}{2d} \left[1 - \left(\frac{R_d}{Pd} \cos \theta \right)^2 \right], \text{ and} \\ FTF &= \frac{F_r}{2} \left(1 - \frac{R_d}{Pd} \cos \theta \right), \end{aligned} \quad (2.3)$$

In Eq. (2.3), these defect frequencies depend on the following parameters: the number of rollers

(n), the shaft speed in hertz (F_r), the contact angle (θ), the roller diameter (R_d), and the pitch diameter, (P_d).

Step-2. As the envelope power spectrum reveals the bearing defect frequency peak and their harmonics in Step-1, Gaussian mixture model-based windows (GMM_{window}) are then created around the defect peaks and the harmonics of each defect frequency to obtain the residual components of the envelope power spectrum, as can be seen in Step-2 of Figure 2.5. The coefficients of the $GMM_{window}(k, \eta)$ are defined as follows:

$$GMM_{window}(k, \eta) = \sum_{i=1}^n \exp \left(-\frac{1}{2} \left(\eta \frac{(k - S_i)^2}{\frac{N_{rfreq}}{2}} \right) \right) \text{ and, } S_i - f_{range} \leq k \leq S_i + f_{range} \quad (2.4)$$

where S_i defines the i^{th} harmonic of the defect frequency and n represents the number of harmonics (e.g., $n=3$ in this study), which is used to calculate the DDR. N_{rfreq} are the number of frequency bins in the range $S_i - f_{range} \leq k \leq S_i + f_{range}$ as below:

$$N_{rfreq} = 2 \cdot f_{range} / f_{resu.} \quad (2.5)$$

f_{range} maintains frequency range in DDR calculation. The parameter of N_{wfreq} can be calculated by measuring the DDR as below:

$$N_{wfreq} = \left(\left(2 \times \frac{1}{100} \right) \times S \right) / f_{resu.}, \quad (2.6)$$

s.t., $span_S = \left(\frac{2}{100} \right) \times S.$

where $f_{resu.}$ defines the resolution of window size ($f_{resu.} = 0.1$ Hz in this chapter). Likewise, η represents a Gaussian random variable that is inversely proportional to the standard deviation, which can be computed as follows:

$$\eta = \left(N_{rfreq} / N_{wfreq} \right) \sqrt{2 \ln p}. \quad (2.7)$$

N_{wfreq} defines the number of frequency bins around the defect frequency components and a fixed value of p in the range $0 < p < 1$, ($p = 0.1$)

Step-3. As the Gaussian window is defined in Step-2, the defect frequency components are calculated by multiplying the Gaussian window, $GMM_{window}(k, \eta)$, around the BPFO, BPF1, or $2 \times$ BPF and harmonics in the obtained envelope power spectrum.

Step-4. The residual frequency components are defined by subtracting the defect frequencies component (in Step-3). As the defect frequency and residual component are defined, the DDR is calculated as the ratio of the defect frequencies and the residue frequencies as below:

$$DDR = 10 \cdot \log \left(\sum_{n=1}^3 \left\{ \frac{\sum_{j=1}^{N_{wfreq}} D_{n,j}^2}{\sum_{j=1}^{N_{rfreq}} R_{n,j}^2} \right\} \right) (dB). \quad (2.8)$$

Here, $D_{n,j}$ and $R_{n,j}$ are the magnitudes of the j^{th} frequency bin for the defect frequency components and residual frequency components, respectively, around the n^{th} harmonic of the defect frequency.

Additionally, to obtain an informative, efficient, and useful description of the DDR values, it is necessary to choose a good mother wavelet function that adequately describes the signal and the terminal nodes (or a set of wavelet coefficients) resulting from the wavelet decomposition process. Thus, this chapter explores the effect of the wavelet filter such as Daubechies mother wavelet function family (db1 to db45), which is one of the mother wavelet function families that is most widely used for analyzing the diverse symptoms of bearing failures [43][1]. To do this, this chapter utilizes a modified 2-D visualization describing the DDR (e.g., degree of defectiveness ratio) relationship among all bearing failures (e.g., a ball pass frequency for the outer raceway (BPFO), a ball pass frequency for the inner raceway (BPF1), twice ball spin frequency ($2 \times$ BSF)), which can be defined as follows:

$$DDR_{overall} = \frac{DDR_{avg}}{DDR_{diff}} \quad (2.9)$$

Where,

$$DDR_{avg} = \frac{DDR_{BPFO} + DDR_{BPFI} + DDR_{2 \times BSF}}{3} \quad (2.10)$$

and,

$$DDR_{diff} = \frac{\left(|DDR_{BPFI} - DDR_{BPFO}| + |DDR_{BPFI} - DDR_{2 \times BSF}| + |DDR_{BPFO} - DDR_{2 \times BSF}| \right)}{3} \quad (2.11)$$

In Eq. (2.9), $DDR_{overall}$ is the ratio of DDR_{avg} and DDR_{diff} such that higher values of $DDR_{overall}$ indicate higher DDR values for all bearing failures or *vice versa*, in which DDR_{avg} is an average value of the DDR for all bearing faults, and DDR_{diff} is an average value of the DDR differences among all bearing defects. This approach takes into account a mother wavelet function and terminal nodes (*i.e.*, WPT sub-bands) that provide low DDR_{diff} and high DDR_{avg} for the purpose of an accurate DDR .

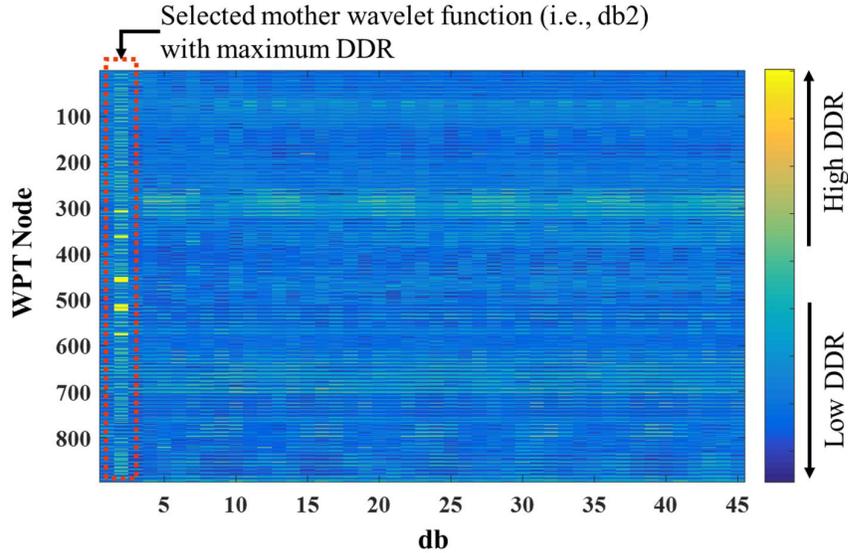


Figure 2.6: Two-dimensional visualization tool of the percentage of each $DDR_{overall}$ for selecting the wavelet filter

Figure 2.6 shows the modified 2-D visualization of the percentage of each $DDR_{overall}$; high $DDR_{overall}$ values are primarily achieved in the high-frequency terminal nodes. Furthermore, larger values of $DDR_{overall}$ are generally yielded when db2 is used as a mother wavelet function than when using other mother wavelet functions. Hence, this chapter used db2 for final DDR calculation about bearing health state representation that is resulting from the seven-level WPT.

Figure 2.7 presents a 2D visualization of DDR values for BCO fault (see Table 1), which is calculated using db2 mother wavelet function in the WPT decomposition. According to the result in Figure 2.7, the proposed 2D visualization tool is highly efficient to visualize bearing outer fault health state.

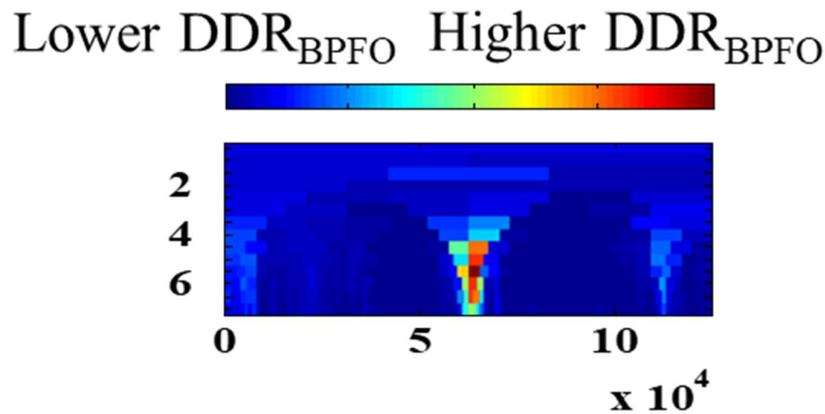


Figure 2.7: An example of 2D visualization of DDR values for BFO signal.

2.3 Results and Discussions

In this section, this chapter investigates the influence of degree-of-defectiveness ratio (DDR) for selecting informative sub-band signal to ensure robust condition-based monitoring of bearing

failures. To verify the results, this chapter uses three fault types (e.g., BCO, BCI, and BCR) in different rotational speeds (see Table 2.2). Figure 2.8 displays original time-domain signals of BCO, BCI, and BCR respectively.

As explained in Section 2.1 that bearing defect frequency and its harmonics are rarely found in the frequency spectrum of the original signal using a traditional signal processing technique, such as fast Fourier transform (FFT) [3]. To explain this problem, in Figure 2.9, this study applies FFT to a real AE signal for the outer raceway (BCO) fault case. According to the result is shown in Figure 2.9, BPFO (i.e. 26 Hz for 300 RPM in this chapter) and its harmonics are not easily

observable in the frequency spectrum (i.e., Figure 2.9(c)) because the high-frequency carrier signal distorts intrinsic information about a BCO.

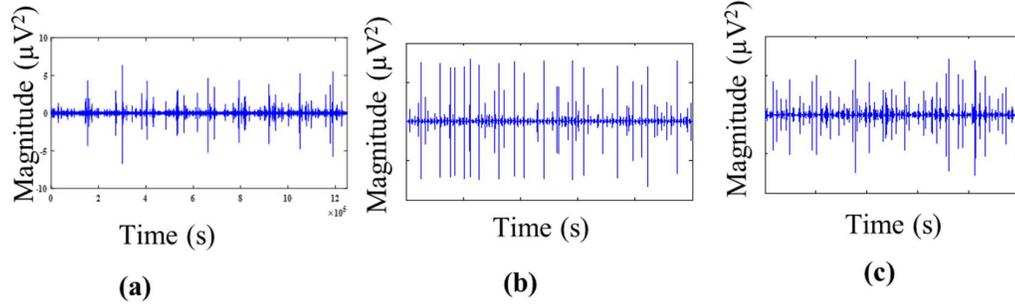


Figure 2.8: Original AE signals of each bearing at speed conditions (a) BCO at 300 RPM, (b) BCI at 350 RPM and (c) BCR at 400 RPM

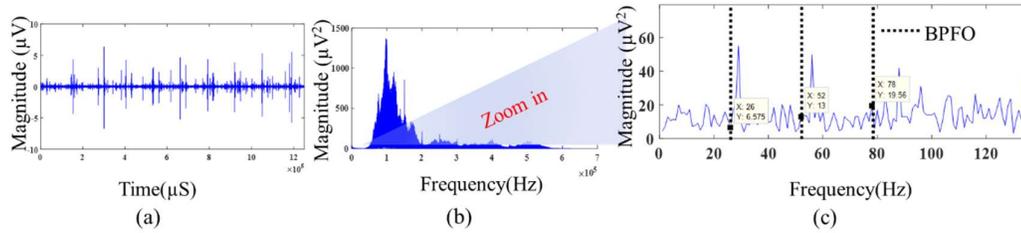


Figure 2.9: The results of FFT and its limitation for finding the defect peak. (a) An original AE signal of the BCO fault, (b) FFT of the signal in (a), and (c) enhanced view of (b) where there is no BPFO or its harmonics.

Though kurtogram analysis is widely used for finding informative sub-bands regarding abnormal fault symptoms, it is still important to have an appropriate degree of defectiveness measure. This chapter, therefore, improves spectral kurtosis value (SKV) based sub-band analysis in [38] by developing a new evaluation metric of the DDR for the proposed WPT-EA. Figure 2.10 compares the result between the proposed WPT-EA with DDR in Figure 2.8 (b) and SKV in Figure 2.10 (a). According to the figure, it is clear that the proposed evaluation metric is highly efficient for finding the three informative sub-band signals of BPFO, BPFI, and $2 \times$ BSF for the outer, inner, and roller raceway faults, respectively. Another important point to note is that SKV based analysis is incapable of selecting informative sub-band information since it misses the defect frequencies, BPFO, BPFI, $2 \times$ BSF, as well as their harmonics in the corresponding sub-bands spectrum views (i.e., the right of Figure 2.10 (a)), while the proposed WTP-EA with DDR method is highly capable of finding appropriate sub-bands, as can be seen in the spectrum views (i.e., the right of Figure 2.10 (b)).

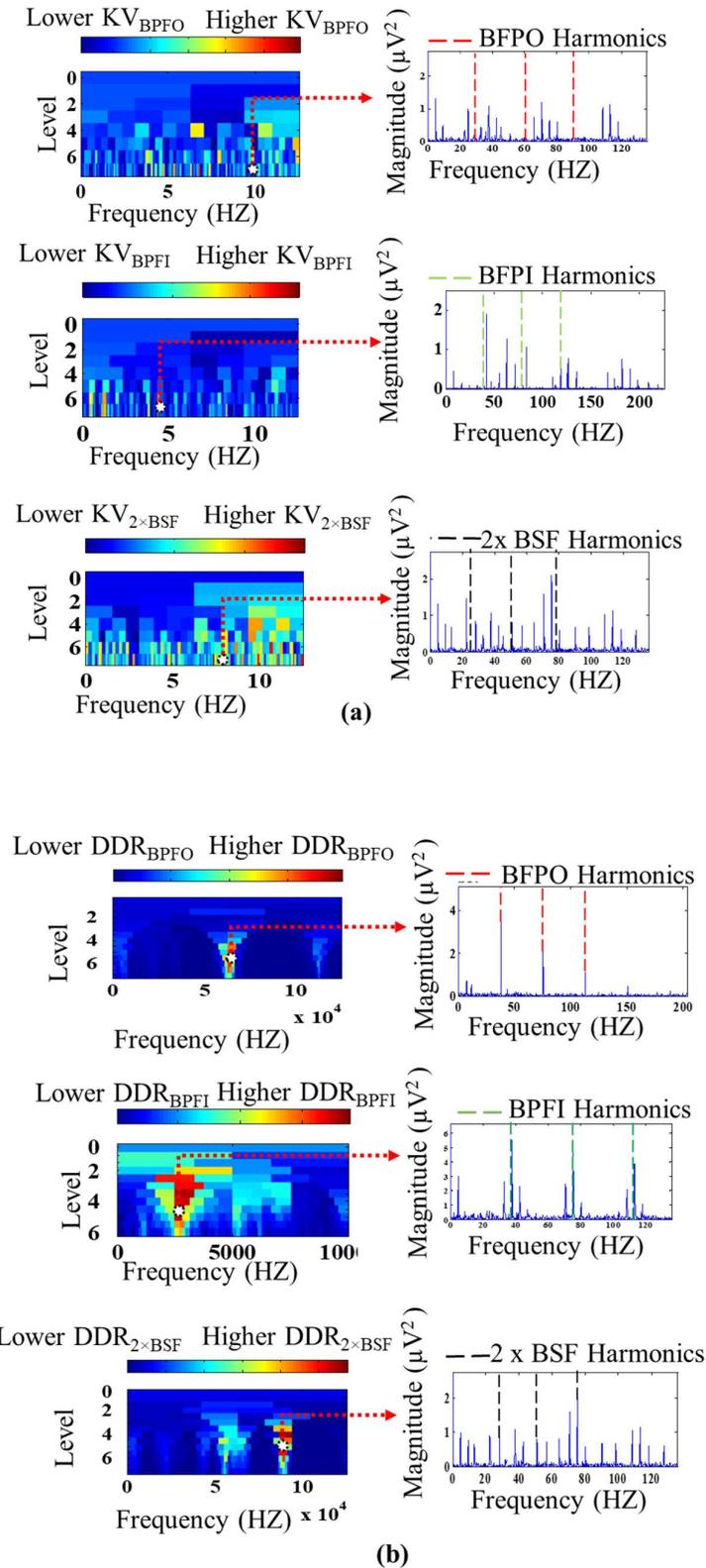


Figure 2.10: 2D visualization tool for finding informative sub-band signals based on (a) SKV and (b) the proposed DDR values

Additionally, this chapter verifies the proposed method for detecting different defects (such as BCO, BCI, and BCR) at the same RPMs. Figure 2.11 presents the time-domain signals of BCO, BCI and BCR faults at 300 RPM and 400 RPM. Figure 2.12 present the result proposed (WPT-EA with DDR) method for identifying bearing defect at the same RPM. According to the result, it is clearly seen the proposed method is also highly effective for the same RPM fault types.

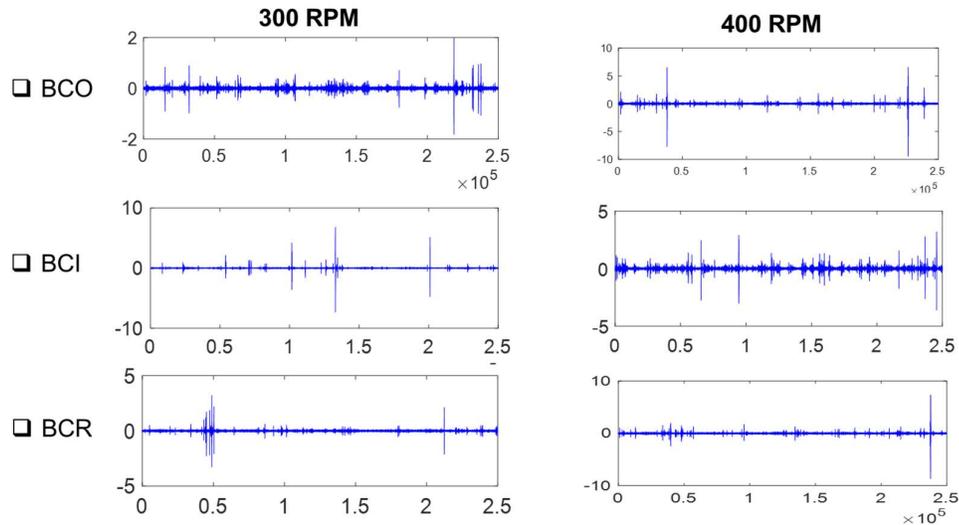


Figure 2.11: Original AE signals of different faults such as BCO, BCI, and BCR at 300 RPM (Left) and at 400 RPM (Right)

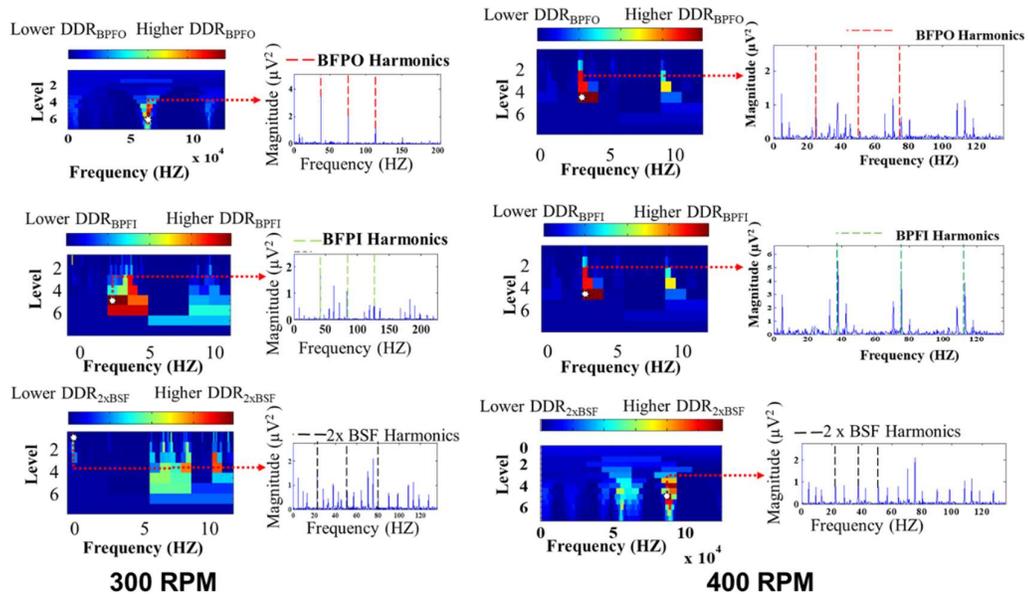


Figure 2.12: Effectiveness of the proposed method for detecting different bearing defects for (a) 300 RPM (left) and (b) 400 RPM (Right)

2.4 Conclusion

In Chapter 2, this dissertation concentrated on early identification of the presence of various bearing failures through WPT-based envelope analysis with DDR evaluation metrics. This study applies WPT to generate a series of sub-band notes and then DDR metrics are applied on each sub-band to quantify. A 2D tool is developed to visualize these DDR values to select the most informative sub-band in which bearing defect frequencies are detectable. The experimental results showed that the proposed condition monitoring scheme outperforms other conventional condition monitoring methods.

Chapter 3

Reliable Multi-Fault Diagnosis Scheme based on Heterogenous Feature Model and Hybrid Feature Analysis

3.1 Introduction

Bearings are crucial mechanical apparatuses in the rotating machinery. They are also leading causes of failure in essential industrial equipment, such as rolling machines in steel mills, paper mills, and wind-turbine power generation plants, where bearing faults account for 51% of all failures [25, 46-48]. The fault diagnosis of bearings is necessary to keep the machines operating normally, reduce abrupt breakdown time and elevated repair cost, and increase productivity [25]. Such challenging scenarios have prompted extensive research into fault diagnosis by modeling essential information about diverse bearing failures [49].

The model-based fault detection and diagnosis (FDD) technique relies on appropriate dynamic models of the systems and has been widely employed to enhance the reliability and availability of industrial systems subject to faults [47, 50]. Today's industrial processes, however, have become increasingly complicated, allowing less tolerance for performance deterioration and productivity decrease [51]. This suggests that the conventional model-based FDD scheme is impractical and often naïve in defining an appropriate physical model for such complex machinery. Fortunately, the rapid development of data mining, data acquisition techniques, and machine learning techniques since 1990 has provided the ability to collect and store an immense amount of process data to extract useful information inherent in such a highly significant amount of recorded data (e.g., acoustic emission (AE), vibration, and current) for the purpose of reliable fault diagnosis in major important industries [4, 11, 51, 52]. It is, therefore, necessary to develop fault diagnosis methods that can

efficiently process a massive amount of data to self-learn fault features and intelligently obtain accurate diagnosis results. Vibration signals and motor current analysis have been widely exploited in the field of bearing fault diagnosis. These methods mainly established high reliability for high-speed bearings fault diagnosis, which are ranging from a few hundred to a few thousand revolutions-per-minute (RPM), due to the difficulty of capturing intrinsic information about low-speed bearing defects [25, 53, 54]. In contrast with a vibration-based analysis, this study utilizes an AE signals-based diagnosis method since AE techniques are highly effective for early-stage fault detection when machines are operating at low speeds. Data-driven FDD methods using AE signal are often the best suited in the context of industrial applications [25, 55]. Essentially, the data-driven FDD scheme includes data acquisition from the bearing during operation, feature extraction and analysis, and decision-making based on selected features for fault classification [25, 55] [56].

To realize highly reliable data-driven FDD methods, it is necessary to deploy different signal processing techniques to extract intrinsic information about bearing defects from the signal of defective bearings [3, 25, 57]. This chapter applies several signal processing techniques to detect intrinsic information about cracks from signals attained from the bearings. These techniques comprise calculating different features in either the time domain, frequency domain, or time-frequency domain. A feature extraction process based on a single method may overlook the discriminative properties of crack conditions [20, 58, 59]. One of the major contributions of this chapter is constructing a heterogeneous feature extraction model (HFEM) consisting of three simultaneous feature extraction schemes – such as time domain, frequency domain, and time-frequency domain signal analysis.

In addition to time- and frequency-domain features, time-frequency analysis (TFA)-based features are also appropriate for bearing fault classification since TFA simultaneously analyzes a measured signal concerning the fault in both the time and frequency domains where impulse information is detectable [25, 58]. One of the most widely used TFA methods for bearing fault signal analysis is called the wavelet transform. The obvious advantage of wavelet-based signal processing techniques is that they have a good time-frequency localization, which allows for the detection of transients that appear in the signals. Thus, this chapter employs the wavelet pack transform (WPT), which is highly effective at decomposing the signal into mid- and high-band frequency nodes so that defect information can be observable. For this reason, this chapter extracts energy information in the mid- and high-band frequency nodes of WPT.

Thus, the HFEM is significantly a high-dimensional feature vector involving as much information as possible about the process conditions that ensure the availability of all necessary information concerning accurate identification [4]. However, there is still an issue of selecting the

most discriminant features that contain essential information about the mechanical crack being investigated. In practice, however, high-dimensional feature vectors involving a large number of features are either irrelevant or redundant to the aforementioned predictive models (supervised and unsupervised classifiers). These irrelevant or redundant features can be a reason for degraded classification performance for a modern crack identification technique.

To alleviate this issue, discriminatory feature selection is an indispensable part of the pressure vessel crack identification method, and the main goal of this feature selection is to attain a refined subset of a discriminatory feature from the original high-dimensional feature vectors. Recent intelligent FDD techniques have adopted hybrid feature selection technique (HFS) that exploits the advantages of the filter and wrapper methods [60, 61] to ensure best features so that they can improve classification accuracy and reduce the computational complexity in the classification process. The wrapper approach selects the feature subset with feature variables showing the highest classification accuracy for a particular classifier, while the filter approach creates a rank of the feature variables or feature subsets using some property values or an evaluation model [61]. This chapter explores an HFS scheme that combines both the wrapper and filter approaches [57, 60-62].

The feature subsets can be generated by performing a complete, sequential or heuristic search (e.g. GA [12]) of the feature space. A complete search ensures a high-quality feature subset, but it is very costly in terms of computational time. In contrast, a sequential forward floating search (SFFS) [63, 64], which is a variant of sequential forward search, is comparatively faster and provides a good tradeoff between computational complexity and quality of selected optimal features. Since one of the most significant tasks in the HFS scheme is to accurately evaluate feature subsets, this chapter employs improved class separability criteria, a key contribution of this study, —the ratio between local compactness and global separability (LCGS)— for evaluating feature subsets by analyzing class samples distribution in a feature space. Several feature evaluation methods have been proposed depending upon classification accuracy or Euclidean distance-based feature distribution criteria [3, 63, 65]. Kang *et al.* recently proposed a feature subset evaluation method using intra-class compactness and the inter-class distance calculated using average pairwise Euclidian distances [25]. However, they did not consider all possible feature distributions. Moreover, the intra-class compactness value considers dense areas only and ignores samples that are located in less dense areas or on the outskirts of a class, affecting the multiclass distribution. Similarly, a high distance value between two classes can dominate the distance values between other classes, and hence the overall interclass separation value.

To address this limitation of conventional average distance-based methods, this chapter proposes HFS-LCGS that uses the new feature evaluation metric, LCGS, as an objective function for SFFS while feature subsets are evaluated in the filter phase to select a series of suboptimal feature subsets. In the wrapper phase of the HFS-LCGS, these suboptimal feature subsets are further evaluated, by estimating the classification accuracy of SVM classifier [66] to select the most discriminant features. Finally, the selected discriminant feature vector is tested for a low-speed rolling element bearing fault diagnosis application.

The remaining parts of this chapter are designed as follows. Section 3.2 explains the experiment setup and the AE signal acquisition technique. Section 3.3 describes the proposed fault diagnosis scheme including heterogenous feature extraction models (HFEM) and hybrid feature selection (HFS) methodology. Experimental results are given in Section 3.4, and concluding remarks are in Section 3.5.

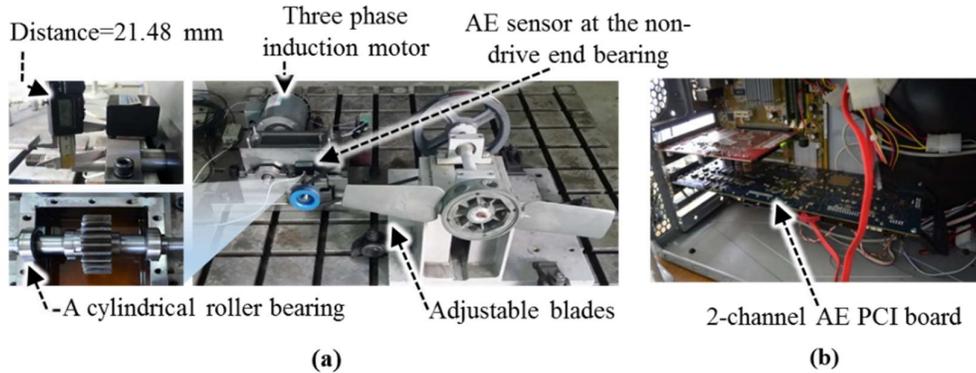


Figure 3.1: Screenshot of the self-designed experiment setup, (a) standard equipment setup, (b) PCI based AE system for data acquisition.

Table 3.1. Summary of acoustic emission (AE) data acquisition conditions, including the use of two different operating conditions and two crack sizes

Dataset	Average Rotational speed (RPM)	Sizes of cracks in the bearing's outer and/or inner roller raceways		
		Length	Width	Depth
Dataset 1 ^a	300	3 mm	0.35 mm	0.30 mm
Dataset 2 ^a	500			
Dataset 3 ^a	300	12 mm	0.49 mm	0.50 mm
Dataset 4 ^a	500			

^a 90 AE signals for each fault type; sampling frequency $f_s = 250$ kHz; each signal is 10 s long.

3.2 Experiment Setup and Acoustic Emission (AE) Signal Acquisition

The standard scheme for measuring the AE signal is introduced in Figure 3.1. We employ some of the most widely used sensors and equipment in the real industries. To capture intrinsic

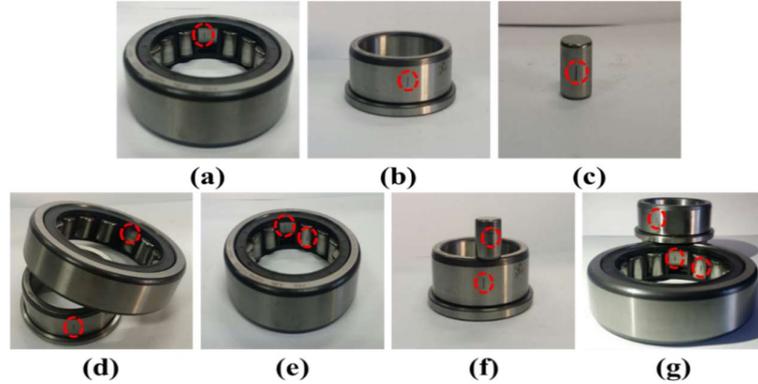


Figure 3.2: Different single and combined bearing faults with 3 mm crack: a) BCO, b) BCI, c) BCR, d) BCFIO, e) BCOR, f) BCIR, and g) BCIOR

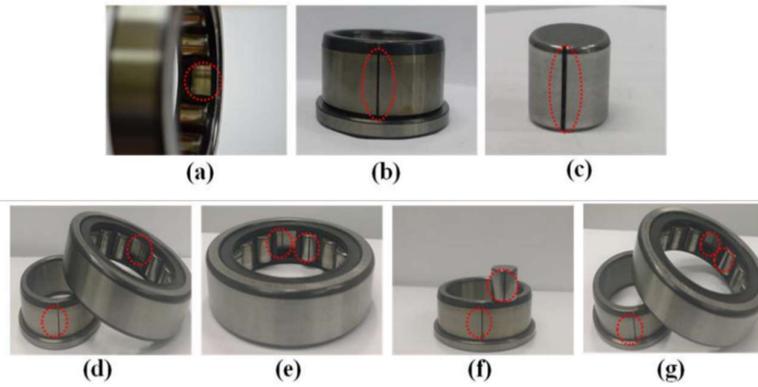


Figure 3.3: Different single and combined bearing faults with 12 mm crack: a) BCO, b) BCI, c) BCR, d) BCFIO, e) BCOR, f) BCIR, and g) BCIOR

information about defect-bearing and bearing with no defect (BND) conditions, the study records AE signals at 250 kHz sampling rate using a PCI-2 system that is connected with a wide-band AE sensor (WS α is from Acoustics Corporation of Physical [3]). The effectiveness of experiment setup and datasets can be studied further in [3, 67]. In this study, AE signals are collected for formulating four experimental datasets of different crack sizes (i.e., small crack and big crack) and different operating speed (i.e., 300rpm, 500rpm). Table 3.1 presents a summary of different datasets. Each

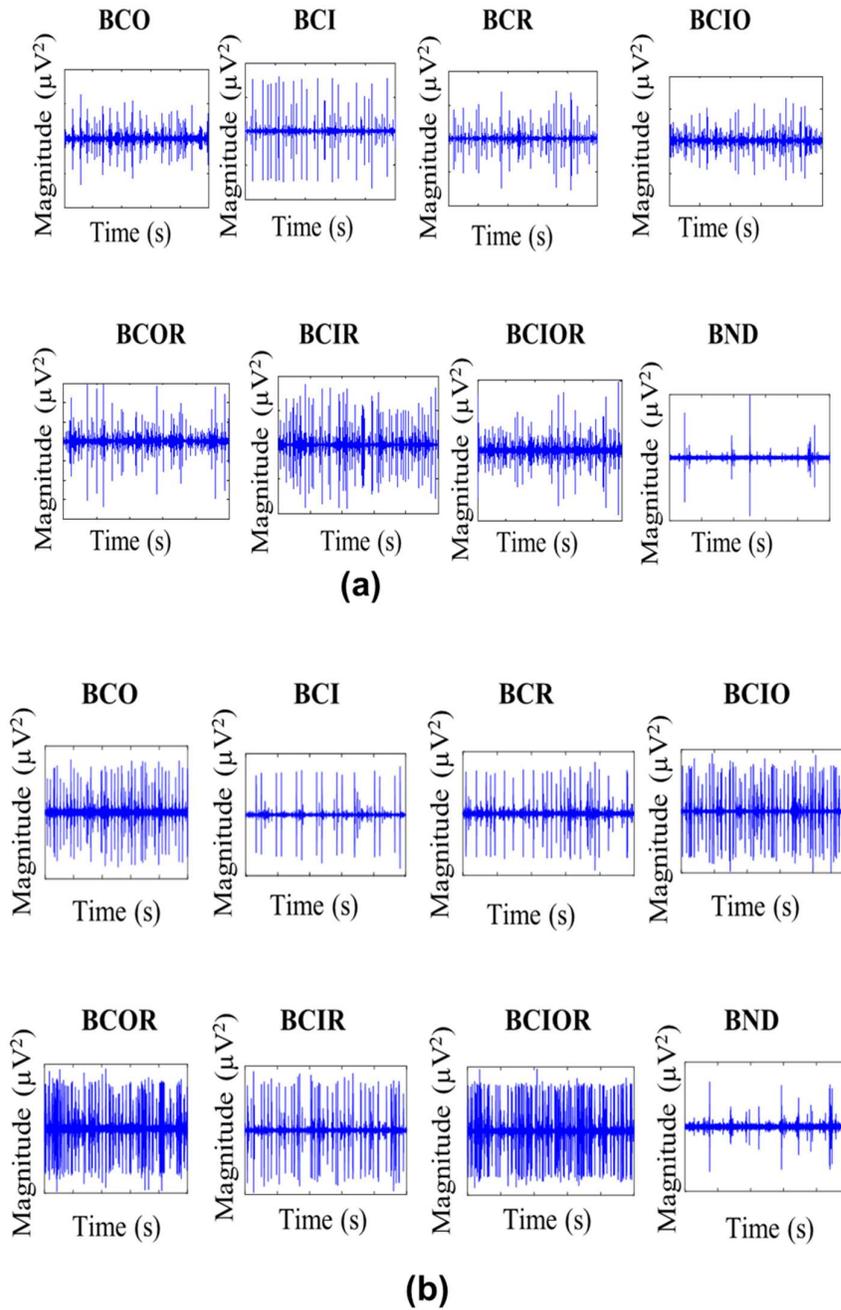


Figure 3.4: Acquired original AE signals of different bearing conditions in (a) dataset 1 and (b) dataset 2.

dataset contains eight types of signal including defect free and seven defective bearings based on crack position: (a) bearing with outer race crack(BCO), (b) bearing with inner crack(BCI), (c) bearing with roller race crack(BCR), and combination of these faults, i.e. (d) bearing with inner and outer cracks(BCIO); (e) bearing with inner and outer cracks(BCOR) (f) bearing with inner and

roller cracks (BCIR) and (g) inner, outer, and roller cracks(BCIOR) and (h) normal condition (BND). Figure 3.2 displays seven defected bearings with 3 mm crack size and Figure 3.3 displays seven defected bearings with 12 mm crack size. The effectiveness of data acquisition can be seen in Figure 3.4 for dataset 1 (in Figure 3.4 (a)) and dataset 4 (in Figure 3.4 (b)). According to results in Figure 3.4, it is clearly seen that each fault condition is significantly varied depending on fault severities such as RPM and crack sizes.

3.3 Proposed Fault Diagnosis Methodology

Figure 3.5 depicts an overall flow diagram of the fault diagnosis method used in the study. The method is composed of the two important processes: a discriminatory feature selection process for deciding the most discriminatory fault signature subset and an online performance evaluation process for validating the effectiveness of the proposed hybrid feature selection.

3.3.1 Heterogeneous Feature Extraction Models (HFEM)

One of the main ideas of this study is that diverse feature extraction paradigms are simultaneously combined to generate a hybrid pool of features, which has the higher discriminative power of accurately detecting each fault.

As explained in Section 3.1, most of the existing studies only focus on the crack detection problem, and no classifier is used to identify fault types. We, therefore, extract heterogenous features from various signal processing domains, namely the time domain, frequency domain, and

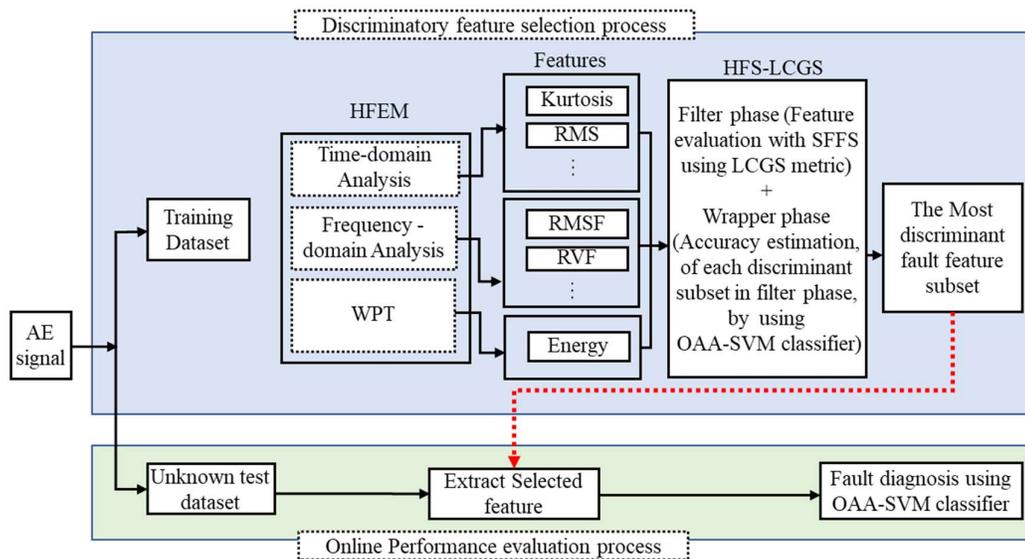


Figure 3.5: An overall block diagram of improved bearing fault diagnosis scheme with HFS scheme

Table 3.2. Definitions of time-domain statistical features used in this chapter

Parameters	Equations	Parameters	Equations	Parameters	Equations
F1	$\sqrt{\frac{1}{N} \sum_{n=1}^N C^2(n)}$	F2	$\frac{1}{N} \sum_{n=1}^N \left(\frac{C(n) - \bar{C}}{\sigma} \right)^4$	F3	$\frac{1}{N} \sum_{n=1}^N \left(\frac{C(n) - \bar{C}}{\sigma} \right)^3$
F4	$\sqrt{\frac{1}{N} \sum_{n=1}^N (C(n) - \bar{C})^2}$	F5	$\frac{\frac{1}{N} \sum_{i=1}^N \left(\frac{C - \bar{C}}{\sigma} \right)^4}{\left(\frac{1}{N} \sum_{i=1}^N C^2 \right)^2}$	F6	$\left(\frac{1}{N} \sum_{i=1}^N C^2 \right)^{1/2} / \frac{1}{N} \sum_{i=1}^N C $
F7	$\max(C) / \frac{1}{N} \sum_{i=1}^N C $	F8	$\max(C) - \min(C)$	F9	$\frac{\max(C(n))}{\frac{1}{N_{\text{tsample}}} \sum_{n=1}^{N_{\text{tsample}}} C(n) }$
F10	$\frac{\max(x(n))}{\text{Standard deviation}}$				

where C is the original time-domain AE signal and there are N samples in the signal

WPT. The main idea of such diversity in the feature extraction process is so that no information about the crack is missed. These features are regarded as discriminative since there is a significant change in the magnitude of the signal when impulses occur due to a crack in the pressure vessel. Therefore, the changes in signal behavior due to a crack can be well-characterized by extracting time-domain statistical feature parameters, such as the root mean square (RMS) (F1), kurtosis (F2), skewness (F3), Standard deviation (F4), Kurtosis factor (F5), Skewness factor (F6), Crest factor (F7), peak-to-peak value (F8), impulse indicator (F9), Margin factor (F10). All the used time-domain features provide statistical properties about the nature of data and were found to be reasonably good features for bearing cracks because they are sensitive to impulse faults [20, 68].

Furthermore, the frequency-domain feature can also discover some important information that cannot be observed in the time domain [20]. Several studies [58, 69] have revealed that the frequency spectrum of the original signals obtained by fast Fourier transform (FFT) provides additional information about the crack, which is helpful for classifying bearing cracks. Thus, the frequency-domain features extracted in this study are as follows: frequency root means square (F11), frequency standard deviation (F12), and mean frequency (F13). Definition of extracted features in the time-domain and frequency-domain are given in Table 3.2 and Table 3.3 respectively.

Table 3.3 Definitions of frequency-domain statistical features used in this chapter

Parameters	Equations	Parameters	Equations	Parameters	Equations
F11	$\sqrt{\frac{1}{N} \sum_{n=1}^N (f(n) - f_s)^2}$	F12	$\frac{1}{N} \sum_{n=1}^N f(n)$	F13	$\sqrt{\frac{1}{N} \sum_{n=1}^N f(n)^2}$

where f is the frequency-domain signal of the corresponding time-domain signal in Table 3.2.

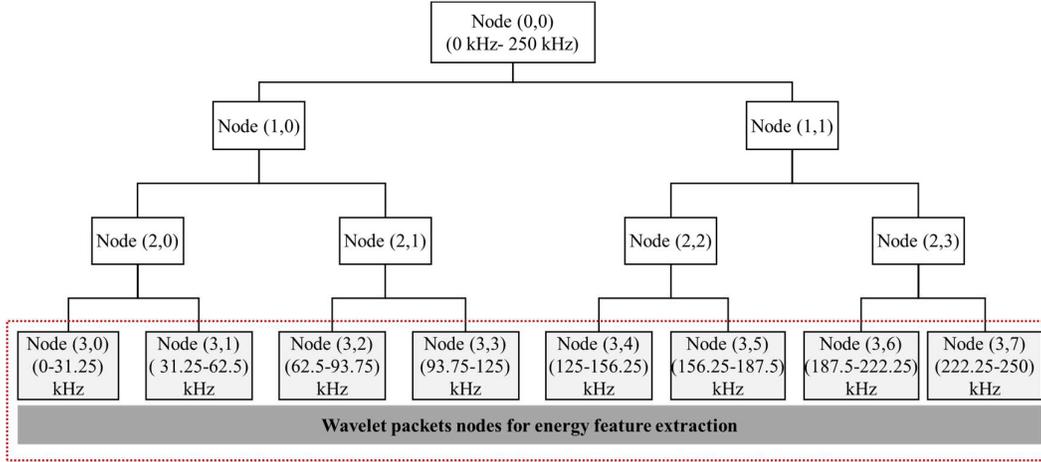


Figure 3.6: The decomposition method of WPT up to three levels for extracting energy features

In addition to time-domain and frequency-domain features, we applied the wavelet pack transform (WPT) pressure vessel signal for obtaining the time-frequency domain features. WPT is highly effective at decomposing the signal into mid- and high-band frequency nodes in which crack information can be observable. For this reason, WPT is applied with a 5 sec AE signal to extract energy information in the mid- and high-band frequency nodes. According to Kang *et al.* [25], the relative wavelet energy (RWE) in the WPT nodes are highly effective for revealing the disordered behavior of the signal due to a crack in the pressure vessel steel. To compute these energy features, we apply a three-level WPT, and we have eight nodes, as shown in Figure 4. Furthermore, the Daubechies 20 (or dB 20) mother wavelet function is used in this study during the WPT decomposing operation. Therefore, RWE can be designed for each node as follows:

$$RWE(k) = \frac{\sum_{i=1}^L w_{k,i}^2}{\sum_{n=1}^{N_{modes}} \sum_{i=1}^L w_{k,i}^2}, \quad (3.1)$$

where N_{nodes} is the total number of WPT nodes (e.g., $N_{nodes} = 8$ in this study), L is the number of wavelet coefficients for each node, and $w_{k,j}$ is the j -th wavelet coefficient of the k -th node.

We extract eight RWE values from WPT nodes, which are denoted as F14-F23. Therefore, we have 23 features, including four time-domain, three frequency-domain, and eight RWE values. In summary, the dimensionality of feature model used in HFS-LCGS feature selection process is $N_{feature} \times N_{anal.samples} \times N_{classes}$ where number features, $N_{feature} = 21$, the number of analysis data samples, $N_{anal.samples} = 30$, and the number of fault classes, $N_{classes} = 8$.

3.3.2 Proposed HFS-LCGS Scheme

Figure 3.7 describes the overall process of the proposed HFS-LCGS, which consists of a filter-based feature selection part in which the feature subsets are evaluated based on local compactness and global separability (LCGS) criteria, and a wrapper-based part that selects the optimal features based on classification accuracy.

It is quite evident that the effectiveness of HFS-LCGS depends on the robustness of the feature evaluation metric, the LCGS when the subsets are assessed with SFFS. As shown in Figure 3.4, this chapter first evaluates the feature subset on randomly selected $1/3^{\text{rd}}$ of the analysis dataset and repeat this process for N iterations. More details of this HFS-LCGS approach are given below.

1) A feature evaluation metric to assess the quality of feature subset in HFS-LCGS

SFFS is used to yield discriminate fault signature subset candidates at the filter-based feature selection phase. To find the useful subset candidates, a precise feature evaluation metric is required. As explained in Section 1, a recent feature evaluation metric based on average Euclidian distance used to measure within-class distance and between-class distance [25]. In practice, the average distance-based feature evaluation metric does not consider the complexity of class and significantly overlooks the overlap in between-class distances. Our study develops an improved evaluation metric as local compactness for within-class distance and global separability for between class distance. To compute local compactness for a specific feature (e.g., f_{nth}), the mean of feature distribution (Cc_M) of each class and the distance of an outmost sample (Cc_dist) (instead of average distance) from the class mean is calculated. Finally, local compactness of each feature variable is calculated by averaging all Cc_dist . Figure 3.8(a) presents the process of local compactness factor calculation and the Eq. (3.2) depicts the final calculation of local compactness (LC) factor.

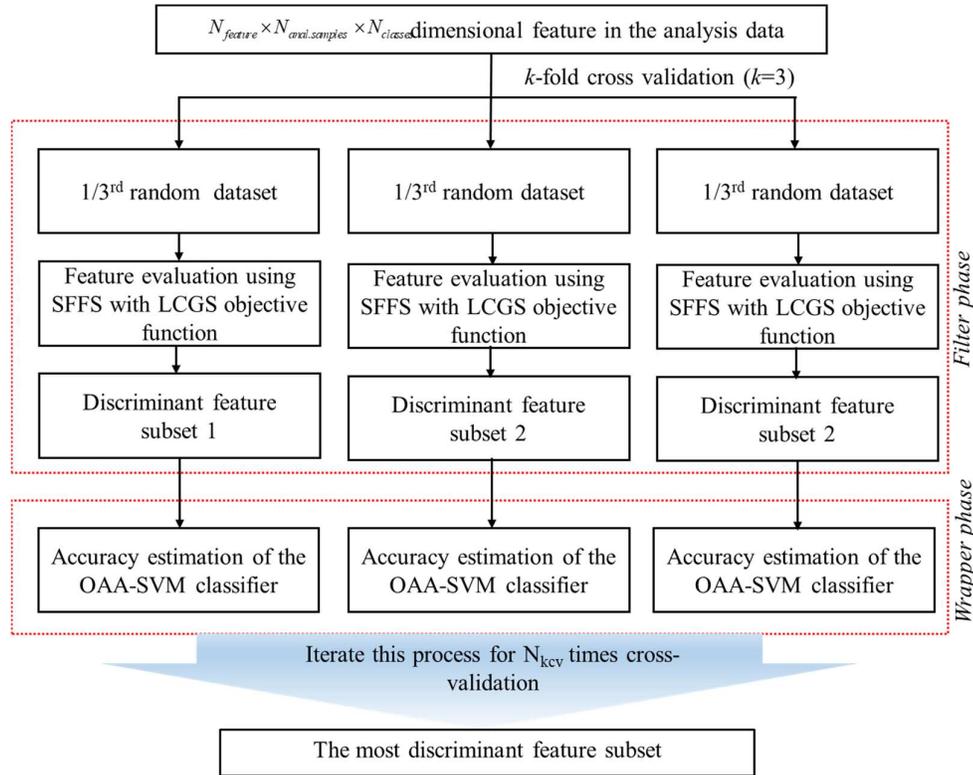


Figure 3.7: Overall procedure of the developed HFS-LCGS scheme for selecting most discriminant feature subset

$$Local_compactness(LC) = \frac{1}{N_c} \sum_{c=1}^{N_c} d_c, \quad (3.2)$$

Where N_c is the total number of classes and d_c is the distance of the outmost sample from the class mean.

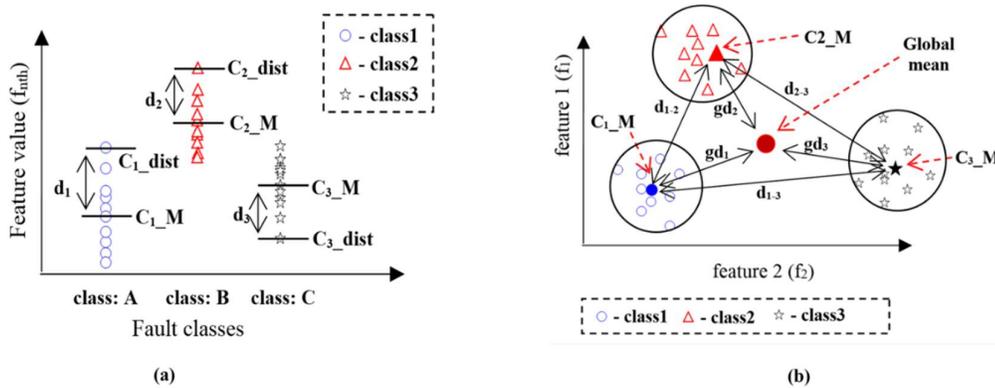


Figure 3.8: The conception of proposed evaluation metric—local compactness and global separability (LCGS) calculation, (a) local compactness (e.g. in 1D feature space) and (b) global separability, where black circles define the local compactness (e.g. in 2D feature space)

To calculate the global separability, a global mean of all samples of all classes and the mean of the individual class are calculated. The class separability distances (i.e. $d_{1,2}$), which represent the distance between the mean of one class to that of another, as well as the global separability distances (i.e. gd_1), which represent the distance between the mean of a class to the global mean, are computed, as can be seen in Figure 3.8 (b) for an example 3-class. Finally, the global separability (GS) is calculated by Eq. (3.3), which can be extended for any finite number of multi-class (i.e. 8 classes in this study) separability measure. The overall process is depicted in Figure .8 (b).

$$Global_separability(GS) = \frac{1}{N_c} \sum_{c=1}^{N_c} \left((gd_c) + \sum_{c \neq j, j=1}^c (d_{c,j}) \right) \quad (3.3)$$

Now that LCGS feature evaluation metric is at hand, we define a function, OBJ, to combine GS and LS, and consider the simplest form in equation (7) to maximize the OBJ function (as. the ratio of the highest value of GS and the lowest value of LC):

$$OBJ = \frac{GS}{LC} \quad (3.4)$$

The evaluation metric in Eq. (3.4) is utilized for any feature subset yield by SFFS to get a series of discriminant feature subset candidates, which are further utilized with classifier in the wrapper approach to obtain the most discriminant features.

2) Accuracy estimation of the SVM classifier in the wrapper method of HFS-LCGS.

In the wrapper-based feature selection of HFS-LCGS scheme, it is necessary to predict classification accuracy of the SVM classifier using a couple of discriminatory feature subset candidates due to the multiple cross-validations in the proposed hybrid feature selection scheme. As seen in Figure 3.6, the execution process for hybrid feature selection is performed using k -fold cross-validation ($k=3$). After N times ($N=10$) cross-validations, this study contains $N \times k$ discriminatory feature subset candidates. Hence, the predictive average classification accuracies are estimated using OAA-SVM classifier for all selected feature subsets for selecting the most discriminant feature.

3.3.3 Fault Classification for Online Diagnosis

As depicted in Figure 3.5, the proposed HFS-LCGS model selects the most discriminant feature elements that are further utilized for online validation of the bearing fault diagnosis model. In the online process, this chapter also uses the one-against-all support vector machine (OAA-SVM) classifier [66] for multi faults classification with linear kernel function to validate our selected feature sets in terms of the classification accuracy. The SVM with linear kernel function classifier is one of the most popular classification methods [70] and is widely used due to simplicity and computational efficiency. SVMs are the most widely used classifier approach and showed significant success in the diagnosis of many real-world applications because of their generalization capability [46].

SVMs are actually designed to solve the binary classification problem by finding the largest margin hyperplane that separates the two classes in the feature space of the training vectors. New data samples are labeled based on the sign of the hyperplane's function, with a positive sign indicating that the new instance of data belongs to the current class and a negative sign indicating otherwise. Let $(x_1, y_1) \dots (x_i, y_i)$ be the training dataset with N samples, where $x_i \in R^D$ is the D -dimensional feature vector, $y_i = \{+1, -1\}$ are the class labels, and $i = 1, 2, \dots, N$. The generalized problem of finding the optimal hyperplane can be expressed in dual form by applying Lagrangian multipliers [18] as shown below:

$$\begin{aligned} \max W(\alpha) &= \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j), \\ \text{subjected } \sum_{i=1}^N \alpha_i y_i &= 0, 0 \leq \alpha_i \leq C, i=1, 2, \dots, N. \end{aligned} \quad (3.6)$$

Here, α_i and α_j are Lagrangian multipliers, x_i and x_j are two input training vectors, and C is a penalty coefficient, which determines the tolerance to misclassification errors and has to be tuned to reflect the noise in the data. The dot product of $\phi(x_i)$ and $\phi(x_j)$ can be replaced with the kernel $K(x_i, x_j)$ by Mercer's theorem [19], [27], as given below:

$$\phi(x_i) \bullet \phi(x_j) = K(x_i, x_j). \quad (3.7)$$

As a binary classifier, SVM cannot be employed directly to solve L -class classification problems. Therefore, the standard OAA-MCSVM scheme is the widely used implementation, where the k^{th} class is trained individually against $L-1$ classes. When OAA-MCSVM finishes its training phase, an unknown test sample x can be classified in Class k for which the decision function Z_i is the highest decisional value in equation (3.8) and the corresponding class label in equation (3.9).

$$Z_i(x) = \sum_i y_i a_i K(x, x_i) + b, \quad (3.8)$$

$$k = \arg \max_{i=1, 2, \dots, L} Z_i(x). \quad (3.9)$$

3.4 Experiment Results and Discussion

The proposed methodology is tested on four datasets with eight fault types obtained from different operating conditions, as shown in Table 3.1. The datasets are divided into two categories: one for offline feature analysis for discriminant feature selection and the other for online evaluation. The analysis datasets consist of 30 of the 90 signals for each fault type for a given speed condition. The remaining 60 signals (is kept higher than analysis data to ensure the reliability of diagnosis performance) of each fault class are used as the unknown signals for online evaluation of the proposed fault diagnosis scheme.

The designed objective function in Eq. (3.4) is highly effective for assessing the quality of the features by calculating LCGS value. As explained in Section 3.2.2, $N \times k$ (or total 30) discriminatory feature subset candidates are created by the proposed HFS-LCGS. Predictive classification accuracy of 30 discriminatory features subject candidates is estimated by OAASVM classifier, as seen in Figure 3.9, Figure 3.10, Figure 3.11 and Figure 3.12 for dataset 1, dataset 2, dataset 3 and dataset 4, respectively. In the wrapper-based analysis, the most discriminatory subset is then selected for each dataset based on not only predictive accuracy but also the frequency of the discriminatory feature subset candidates. To verify the above results of HFS-LCGS, we compare with a state-of-the-art feature evaluation metric using average distance in [25]. Table 3.3 summarizes the most discriminant features for the proposed HFS-LCGS and [25].

Finally, in the online fault diagnosis process, only the most discriminant features (see Table 3.3) are extracted from the unknown signals of the evaluation dataset, and the OAASVM classifier

[66] is applied to calculate the classification accuracy. Additionally, k-fold cross validation (3-cv) [71], an efficient method for estimating generalized classification performance, is deployed to evaluate the diagnostic performance of the

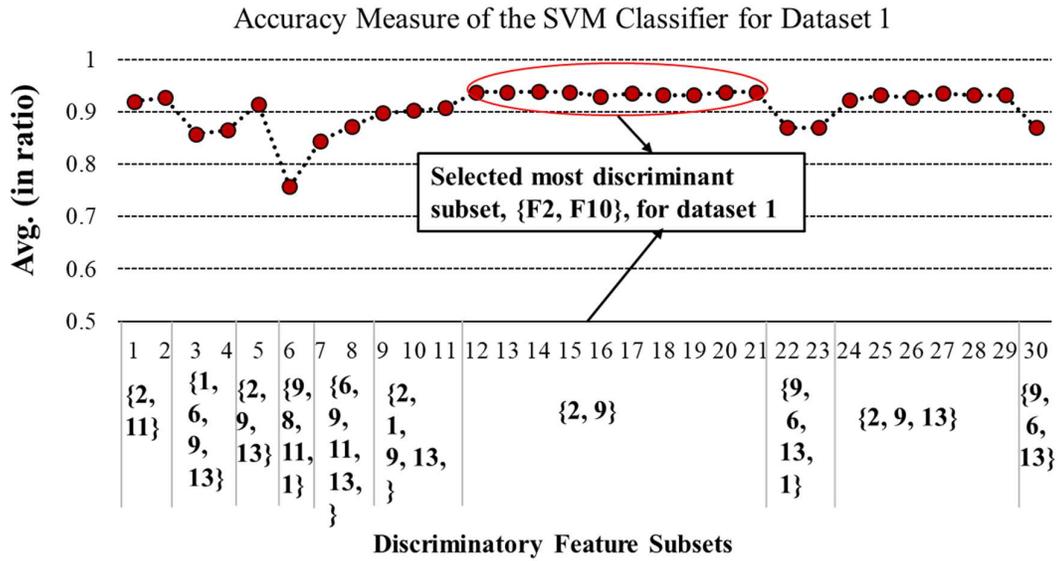


Figure 3.9: Estimated classification accuracy of SVM classifier for 30 feature subsets candidates

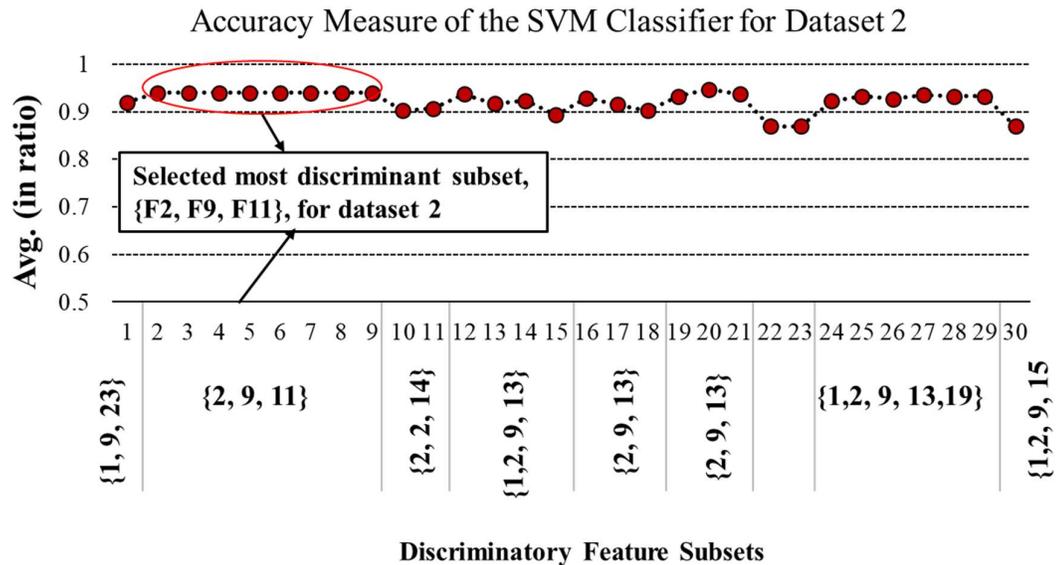


Figure 3.10: Estimated classification accuracy of SVM classifier for 30 feature subsets candidates

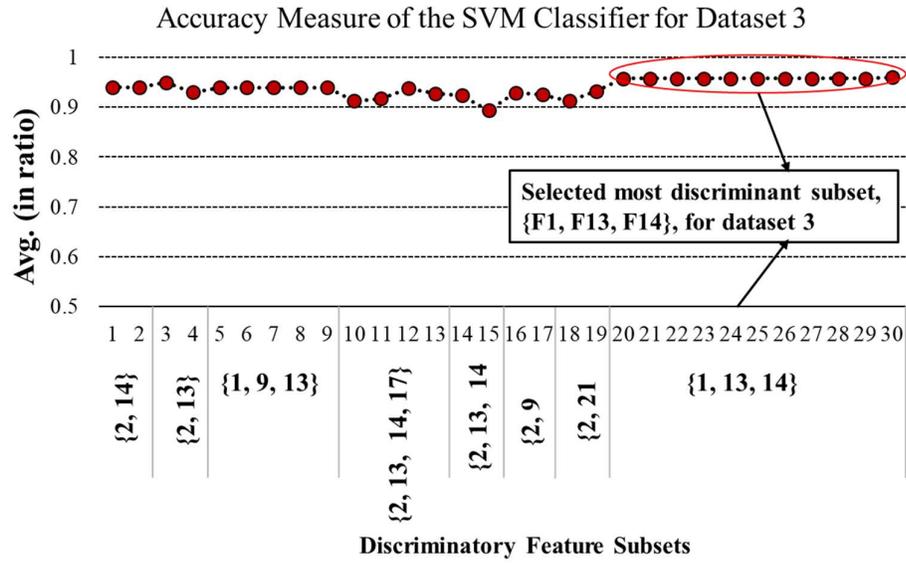


Figure 3.11: Estimated classification accuracy of SVM classifier for 30 feature subsets candidates

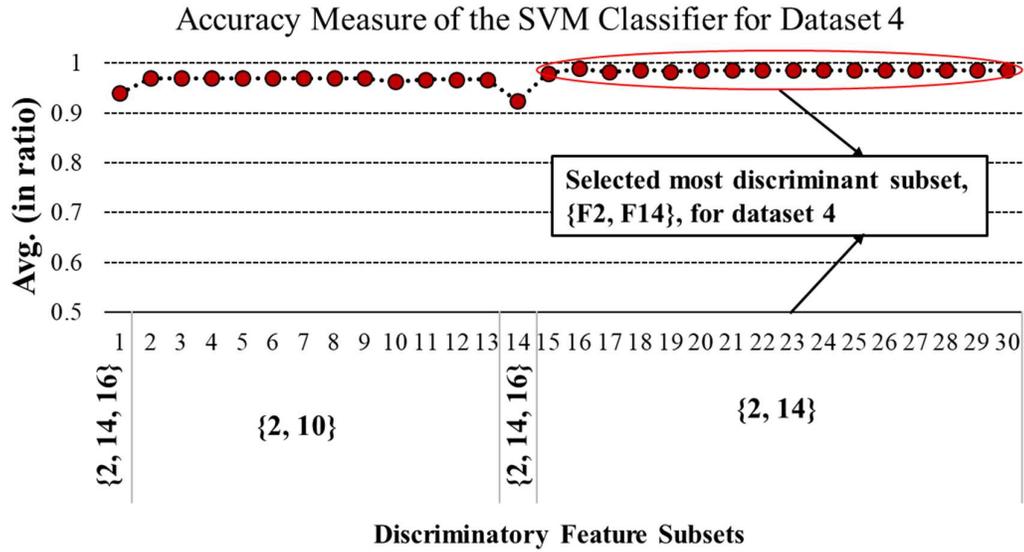


Figure 3.12: Estimated classification accuracy of SVM classifier for 30 feature subsets candidates

proposed HFS-LCGS method versus the state-of-art algorithm [25], in terms of average classification accuracy (avg.) and sensitivity that are defined as below [71]:

$$Sensitivity = \frac{N_{rtp}}{N_{rtp} + N_{fn}} \times 100\% \tag{3.5}$$

$$avg. = \frac{\sum_{k=1}^L N_{rtp}}{N_{sample}} \times 100(\%). \quad (3.6)$$

where sensitivity is the number of positive classes that are correctly identified as positive, L is the number of fault classes or categories (i.e. $L=8$ in this study), N_{rtp} is the rate of true positives, and the number total samples is N_{sample} .

Table 3.3. The most discriminant feature set of four datasets

Datasets	Methodology	
	The most discriminant feature subset by HFS-LCGS (Proposed)	The most discriminant feature subset by state-of-the-art algorithm [25]
Dataset 1	F2, F10	F10, F11, F13
Dataset 2	F2, F 9, F15	F9, F10, F13, F17, F19
Dataset 3	F1, F13, F18	F2, F10, F11, F19, F20}
Dataset 4	F2, F14	F2, F12, F13, F17

Table 3.4. Average sensitivities of the three different models

Datasets	Methodologies	The average sensitivity of each fault types								
		BCO	BCI	BCR	BCIO	BCOR	BCIR	BCIOR	BND	Avg. (%)
Dataset 1	All features	87.57	85.84	77.10	78.37	81.44	86.24	83.00	83.11	82.83
	[4]	85.00	93.00	96.33	93.00	95.00	90.55	91.00	90.43	91.79
	Developed HFS-LCGS	96.74	95.68	96.48	96.74	93.68	96.61	96.88	96.06	96.11
Dataset 2	All features	80.66	82.13	82.66	82.53	82.53	82.39	82.26	82.21	82.17
	[4]	89.28	93.08	94.62	89.48	96.22	91.35	90.22	93.97	92.28
	Developed HFS-LCGS	98.13	98.93	97.47	99.33	99.07	99.20	99.07	98.82	98.75
Dataset 3	All features	91.69	92.00	93.69	92.49	82.09	90.36	91.96	96.00	91.29
	[4]	91.9	95.1	83.03	93.57	98.3	92.77	93.03	95.43	92.89
	Developed HFS-LCGS	100.00	100.00	98.36	100.00	98.00	98.80	100.00	99.40	99.32
Dataset 4	All features	92.31	93.92	93.12	92.22	93.31	92.17	98.00	93.30	93.54
	[4]	100.00	100.00	100.00	100.00	100.00	96.00	92.59	99.50	98.51
	Developed HFS-LCGS	100.00	100.00	100.00	100.00	100.00	99.33	100.00	99.92	99.91

Experimental results listed in Table 3.4 clearly demonstrate that the proposed feature selection model outperforms the other approaches under different conditions. In the datasets for small crack size, the weakly generated fault signals are not significantly distinguishable, affecting the classification performance. The proposed feature selection model selects the best subset of features with the best distribution in the high-dimensional feature space to increase the classification performance of the fault diagnosis system. In contrast, the existing average distance-based approaches do not consider the distribution of features and render a reduced classification performance.

Figure 3.13 presents the confusion matrices of the described method for four datasets (see Table 3.1). As it can be interpreted, the classifier has categorized fault signals of eight classes (BCO, BCI, BCR, BCFO, BCOR, BCIR, and BCIOR and BND). So, based on the outcome, it can be said that the proposed method with hybrid feature selection for fault diagnosis is highly accurate.

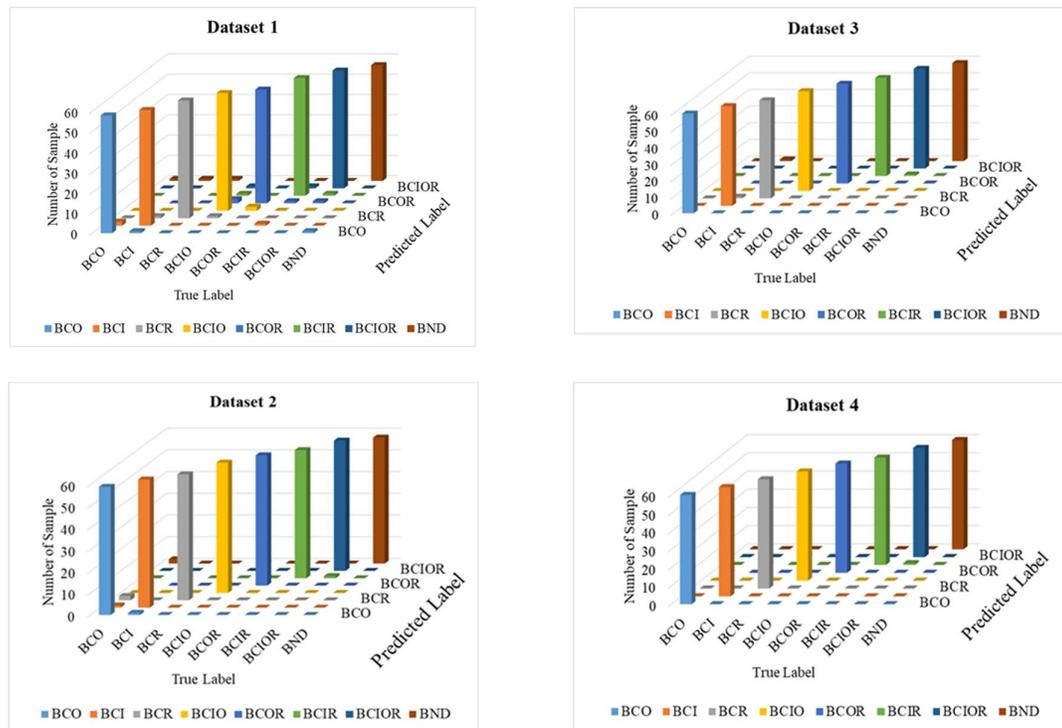


Figure 3.13: The results of confusion metrics of the proposed method to demonstrate actual vs. predicted deviation for four different datasets

In addition to average classification accuracy (avg.) and sensitivity-based performance measurement, this chapter also calculated precision and F1 score to verify the reliability of the

proposed method. Precision or positive predictive value (PPV) is calculated in the following equation,

$$PPV = \frac{TP}{TP + FN} \tag{3.7}$$

where, TP is true positive counts and FN is false negative counts.

F1 score or the harmonic mean of precision and sensitivity is calculated as follows,

$$F1 = 2 \cdot \frac{PPV \cdot TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN} \tag{3.8}$$

Figure 3.14 displays the result of sensitivity and precession of the proposed method. Table 3.5 summarizes the result of PPV and F1 for the proposed method for four datasets. In the results, it shows the precision of the proposed method is also significantly high, which demonstrates the reliability of the proposed method.

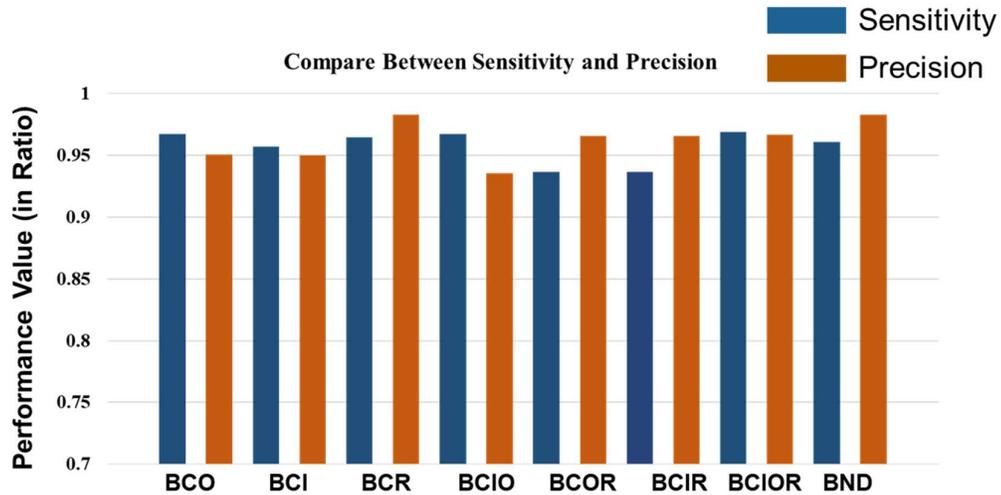


Figure 3.14: Sensitivity and precession results of the proposed method for different fault classes

To further analyze this effectiveness phenomenon, our chapter exploits a two-dimensional representation of discriminative feature selected by the developed HFS-LCGS and by the state-of-the-art method. It is evident from Figure 3.15 that the proposed algorithm selects a feature subset with most separable class distribution compared to its counterpart.

Table 3.5. Proposed Method robustness analysis based on PPV and TRP under various bearing fault conditions for different dataset

Data sets	Perform ance Matrices	Average PPV and TRP of each fault types (in Ratio)								ACA/F1 (Ratio)
		BCO	BCI	BCR	BCIO	BCOR	BCIR	BCIOR	BND	
Data set 1	PPV	0.96	0.95	0.97	0.96	0.93	0.96	0.967	0.95	0.96/0.95
	TRP	0.97	0.956	0.96	0.97	0.93	0.96	0.97	0.96	
Data set 2	PPV	0.921	0.983	1.0	1.0	1.0	1.0	0.983	1.0	0.98/0.98
	TRP	0.983	0.989	0.97	0.99	0.991	0.992	0.990	0.98	
Data set 3	PPV	1.0	0.967	1.0	1.0	1.0	1.0	0.983	1.0	0.99/30.99
	TRP	1.0	1.0	0.98	1.0	0.98	0.98	1.0	0.99	
Data set 4	PPV	1.0	1.0	1.0	1.0	1.0	1.0	0.983	1.0	0.99/0.99
	TRP	1.0	1.0	1.0	1.0	1.0	0.993	1.0	0.999	

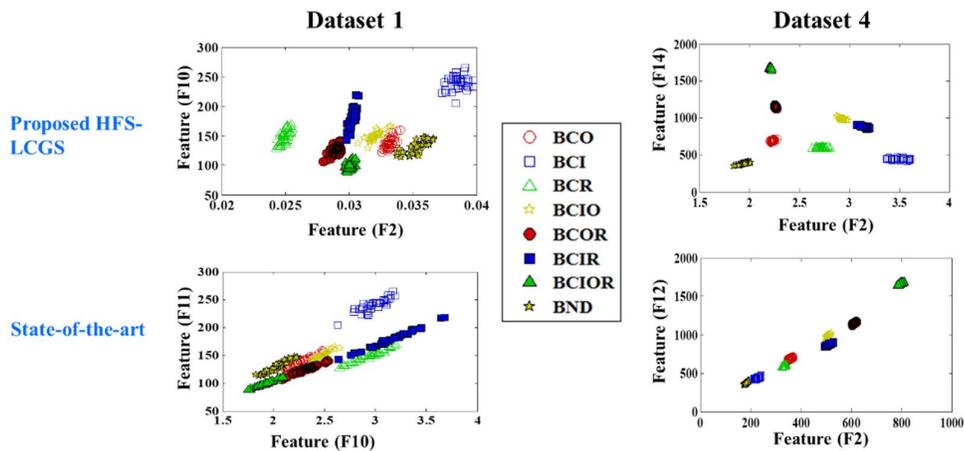


Figure 3.15: 2D visualization of most discriminant feature subset results by developed HFS-LCGS and the state-of-the-art for (a) dataset 1 and (b) dataset 4

3.5 Conclusions

The HFS-LCGS method was developed to select discriminant features by calculating local compactness and global separability (LCGS), which is not achievable in traditional Euclidian distance-based separability. The key contribution of HFS-LCGS is to assess the quality of feature subsets based on LCGS. This evaluation metric is designed as the ratio of local class compactness and global separability. Using this evaluation metric, SFFS yields discriminant feature subset

candidates and the most discriminant feature subsets, which is finally determined via accuracy estimating of the SVM classifier in the wrapper approach. Experimental results indicate that the proposed method is more effective for identifying the most discriminatory feature subset by achieving diagnostic performance improvements from 1.4 % to 17.74% in average classification accuracy. Moreover, this study proves that the proposed hybrid feature selection can effectively reduce computational overhead for fault diagnosis since HFS-LCGS selects the most compact features subset from a high-dimensional features vector

Chapter 4

Acoustic Spectrum Imaging and Deep Convolution Neural Network for Rotating Bearings Fault Diagnosis Under Variable Operating Conditions

4.1 Introduction

Electromechanical motors such as induction motors have widespread applications in wind turbines, pumps, and generators where they are accountable for almost 70% of the gross energy consumption [72, 73]. Bearings are used to minimize friction in motors which undergo rotation. Moderation of friction preserves energy along with supporting a propitious machine lifetime [74]. However, bearings may exhibit surface spall or cracks due to coarse operating environments and cyclic stuffing. Bearings are the most frequent failed components in rotating machines and account for more than 50% of failures [72, 75]. If these faults in bearings are not detected in the early stage, they can lead to unexpected shutdowns, which are unfavorable in terms of cost and production [76-78].

Fault detection of a bearing is accomplished by collecting data from various sources (e.g., vibration acceleration signals, acoustic emission signals, and motor currents), which has been an essential aspect of studies conducted over the last few decades [77, 79, 80]. These studies prove that diagnosis of the bearing can lessen maintenance expenditures by enhancing the reliability of machinery [80-84]. In the field of bearing fault diagnosis, vibration signals [77, 85] and motor current analysis [53, 86] have been widely exploited. Multiple signature analyses of vibrations and

motor currents have also been considered in research to guarantee improved reliability [87]. Although These analyses have provided high reliability, their focuses have been on fault diagnosis of high-speed bearing since it is cumbersome to capture intrinsic information about low-speed bearing defects from weak vibration and current signals. On the other hand, acoustic emission (AE) signals can capture essential diagnostic information from low-energy signals [39, 41, 88-90], which makes AE signals more significant for data-driven diagnosis approaches. This study employs an AE-based fault diagnosis approach for low-speed bearings.

Fault diagnosis of low-speed bearings under variable speed conditions (e.g., revolutions per minutes (RPM)) is still a challenging problem. A traditional data-driven fault diagnosis scheme that applies several processes, for example, handcrafted feature extraction by a signal processing technique and identification of fault types using extracted features. Authors in [90] proposed an AE-based fault diagnosis technique that extracts informative features and then features are further utilized with classical classifiers such as nearest neighbor (NN), nearest mean (NM), self-organizing map (SOM), and backpropagation neural network (BNN) to investigate the diagnostic performance for fixed working conditions. Although this method reports an improved classification accuracy, it cannot be guaranteed the same performance for changing operating conditions. Recent, acoustic signal-based diagnosis methods mostly consider an envelope analysis-based approach by analyzing the peaks at the character frequencies associated with each defect type in the power spectrum of the envelope signal. The combination of wavelet packet transforms (WPT) and envelope analysis has been studied for AE signal-based fault diagnosis to explore useful sub-band signal for extracting feature for low-frequency harmonic components [91]. However, the relationship between the defect frequency and rotational speed of the bearing is not obvious; thus these approaches are inefficient under variable rotational speeds [16, 43, 92]. Likewise, thermal imaging-based approach [93], conventional feature extraction-based methods [25] have their pros and cons, but issues, as mentioned above, are still unresolved, which result in an ineffective diagnosis performance for variable speed conditions. Moreover, it is necessary to have proper domain expertise to select pertinent features [25, 94, 95] for classical machine learning schemes. Unsupervised feature learning mechanism through the layer-wise perceptron, known as a deep neural network (DNN), is studied together with traditional signal processing techniques in a wider aspect because of their capabilities of learning the features automatically from input data [96]. Several deep learning approaches, i.e., deep belief networks (DBF) [97], deep convolutional neural network (DCNN) [98], deep autoencoder[77, 99] are used to learn automatic features from the input to perform machinery fault classifications. However, feature selection from the raw input data directly is still challenging

because of the various natures (e.g., disparate measuring devices, noise, variable speed conditions, variable load conditions, etc.) of the input signals [99].

This chapter, therefore, presents a new approach to the diagnosis of low-speed bearings under variable speed conditions, which addresses two major limitations in existing methods: a) the requirement of domain level expertise for feature extraction and selection under different operational speeds, and b) the requirement of special dynamic algorithms for automation of the feature extraction process. In this chapter, we develop two-dimensional (2D) acoustic spectral images (ASI) instead of one-dimensional (1D) signals to represent bearing health states appropriately. The 2D ASI employed to observe the pattern for different health states to make the feature extraction and selection process simpler for speed invariance conditions. The proposed 2D ASI creates identical patterns for the same type of health, where variable operating speeds do not affect identical patterns for certain types of the health state. Once a 2D ASI representing the bearing health state is obtained, an adaptive deep convolutional neural network architecture (ADCNN), a variant of the LeNet5 [100] architecture, is proposed to automate the feature extraction and optimal feature selection processes by recognizing patterns from the image pixels. Furthermore, this chapter applies an adaptive learning rate for training the ADCNN. It is essential to have an appropriate learning algorithm; otherwise, the solution could converge to a local minimum. The proposed diagnosis methodology (ASI + ADCNN) is validated in a low-speed bearings fault diagnosis application with four faults, including different crack sizes and operating speeds. In this study, after creating the invariance scenario with 2D ASI, an ADCNN based approach was carried out to resolve these challenges.

The main contributions of the proposed model (ASI + ADCNN) can be summarized as follows.

- 1) A novel identical 2D acoustic spectral imaging (ASI)-based pattern formation method for different health types is developed to explore the potential information of AE signals.
- 2) This 2D ASI is further utilized with a deep convolution neural network (DCNN) to mechanize the feature extraction and selection process.
- 3) ADCNN is proposed for fault diagnosis for low-speed bearings with variations of the shaft speed (e.g., RPMs). The proposed method (ASI + ADCNN), including ASI for the RPM independent pattern and ADCNN for fault diagnosis under variable RPMs, was suitably validated with extensive experiments and simulations, which justifies the potential of the proposed methodology over existing approaches in terms of achieving satisfactory theoretical results compared to experiments.

The remaining part of the chapter is organized as follows. Section 4.2 provides details of the methodology, including the AE data acquisition system, acoustic spectral imaging (ASI)-based bearing health state visualization, and deep convolutional neural network (DCNN) with adaptive learning (ADCNN) for bearing fault classification. The experimental results are provided in Section 4.3. Finally, the chapter is concluded in Section 4.4.

4.2 Methodology

In this study, the target of this dissertation was to classify machine faults under variable speed conditions (i.e., RPM) using a deep learning model, such as proposed ADCNN. The block diagram of the proposed ADCNN-based bearing fault diagnosis framework is presented in Figure 4.1. The proposed method consists of three essential steps. After data acquisition, the first step is to obtain new preprocessing technique for the collected acoustic emission (AE) data to transform the 1D signal into a 2D image, which is referred to as acoustic spectrum imaging (ASI). The main advantage of using such pre-processing is that valuable information about how rotating components are distributed within discrete frequencies. The second step is to feed the preprocessed 2D ASI matrix into an ADCNN for optimized deep learning to produce a model for fault diagnosis. The final step is to apply the ADCNN to for classify faults for variable RPMs.

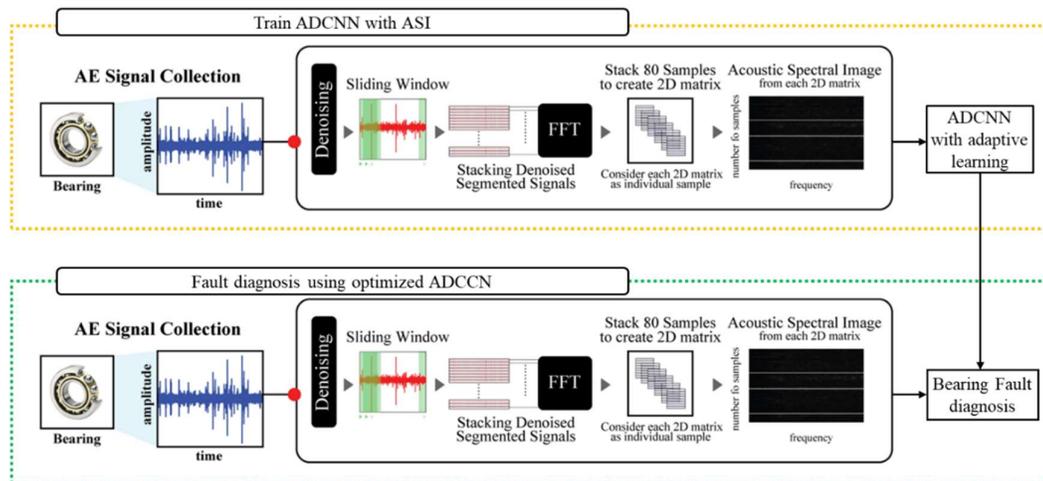


Figure 4.1: The overall structure of the ADCNN-based method for RPM variant fault diagnosis.

4.2.1 Experimental Testbed and Data Acquisition

An experiment was conducted on a self-designed test rig to collect acoustic emissions (AE) signals, as illustrated in Figure 4.2. As the main purpose of this study is to diagnose faults for variable RPMs, the test rig was driven at three different motor speeds of 250, 300, and 350 RPMs.

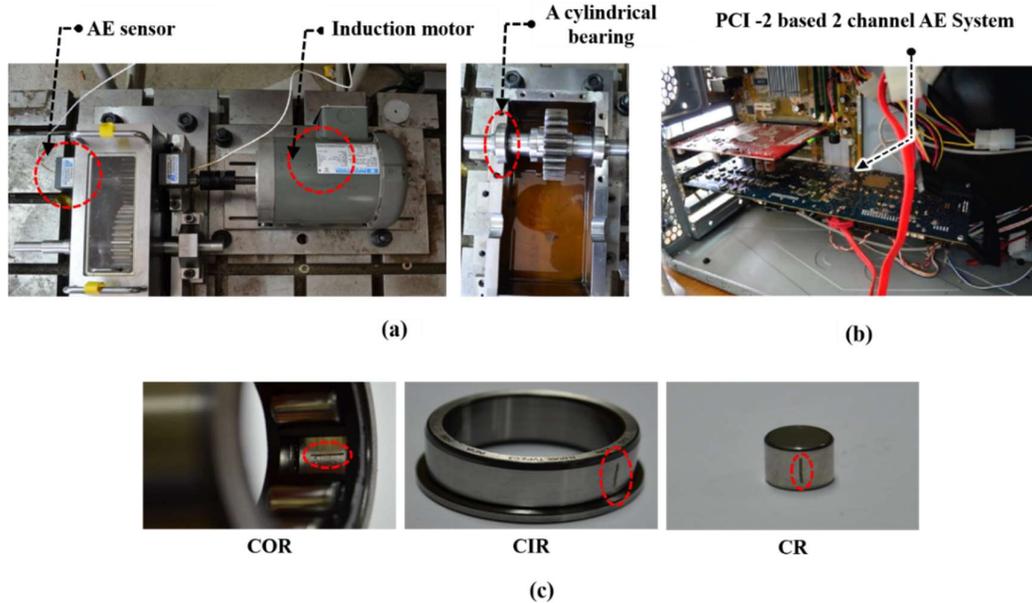


Figure 4.2: (a) Experimental testbed, (b) PCI 2-based data acquisition system, and (c) Examples of ORC, IRC, and RRC bearing defects.

This setup has two shafts, namely a drive end shaft (DES) and a non-drive end shaft (NDES). From this figure, it has been observed that a three-phase induction motor is positioned in the DES at three different speeds (i.e., 250, 300, and 350 revolutions per minutes (rpm)) and the bearing house is connected to the motor shaft through a gearbox with a reduction ratio of 1.52:1. A cylindrical bearing (model FAG-NJ206-E-TVP2) from real industry is used in the experimental testbed. At the NDES, a $WS\alpha$ AE sensor is positioned over the bearing house in the shaft [10]. To quantify the functioning speed, a displacement transducer is installed on the NDES. For capturing the intrinsic information, AE signals were collected at a sampling rate of 250,000 Hz (250 kHz) for 6 mm fault size by using a PCI-2 system. Bearings with three defect conditions of an inner raceway crack (IRC), outer raceway crack (ORC), and roller raceway crack (RRC), as well as one normal condition (NC) bearing, were used to simulate faults for each RPM. Each recorded signal has 1 second of length and in total 110 signals are collected for each health conditions. Detailed specifications of the data acquisition system are given in Table 4.1, and the screenshot of the experimental setup is shown in Figure 4.2(a) that shows the real test rig and Figure 4.2(b) that shows

the PCI-based data acquisition system, which used in this set up to record the AE signal for 2 channels. Additional studies involving the experimental test rig and the data acquisition system can be found elsewhere [4].

Table 4.1. Specification of the proposed data acquisition system

AE sensor (PAC W5a)	Peak sensitivity [V/ μ bar]: -62 dB
	Operating frequency range: 100 to 900 kHz
	Directionality: ± 1.5 dB
	Resonant frequency: 650 kHz
PCI board with 2-channel AE sensor	18-bit 40 MHz A/D conversion
	AE input: 2 channels (one is 10 M samples/s rate, and another is 5 M samples/s as two channels, with simultaneously used)

4.2.2 Acoustic Spectrum Imaging (ASI) for Bearing Health State Visualization

As explained in Section 1, bearing characteristic frequencies are not observable in raw AE signals. Therefore, it is essential to define an appropriate visualization tool that reveals a unique pattern regarding the health state of a bearing. Consequently, this study developed a 2D ASI tool that generates an identical health pattern even with various RPMs for each fault type (e.g., NC, IRC, ORC, and RRC). It is important to provide the full advantages of the 2D structure to the deep neural network with these identical patterns. To carry out the ASI, several steps must be performed. The steps of the proposed ASI are shown in Figure 4, and the detailed process is given below.

The performance of the ASI analysis can be improved if the variations between the flanks and peaks of the spectral energy, which mainly originate from noise from various sources, are minimized. Therefore, this study applied a denoising technique known as pre-whitening on AE signal before ASI [101]. In this chapter, the pre-whitening technique is used to reduce the spectral energy variations in the incoming AE fault signal, as can be seen in Figure 4.3. To obtain the pre-whitened signal, an autoregressive (AR) model, $x(n)$, is utilized, which is defined as follows.

$$x(n) = \sum_{i=1}^k a_i y(i+n) + e(n) \quad (4.1)$$

Here, $y(n)$ is the incoming bearing signal, a_i are the AR coefficients, k denotes the order of the AR model, and $e(n)$ is the residual signal representing a spectrum close to the white noise spectrum.

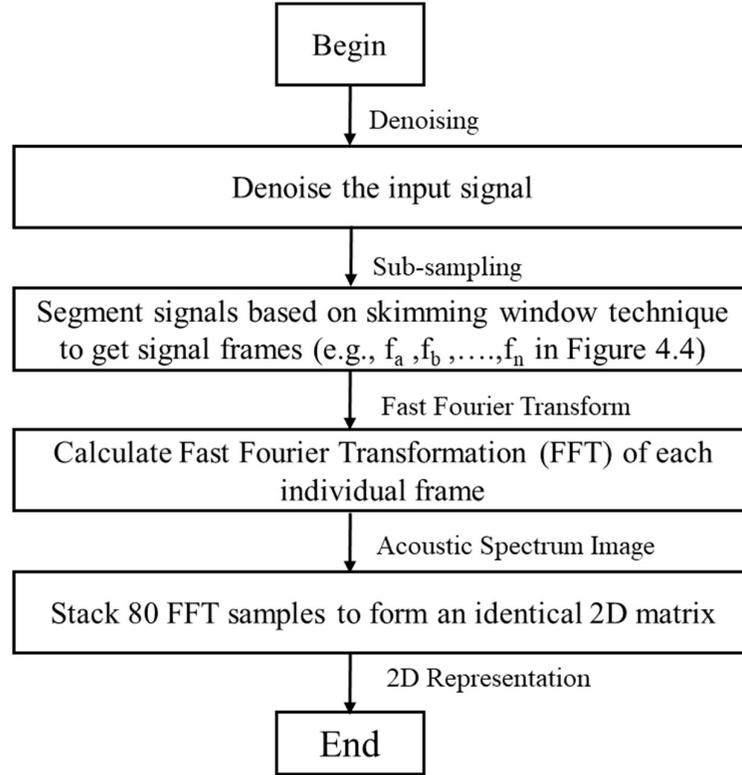


Figure 4.3: Flowchart of the acoustic spectrum imaging (ASI) for visualization of the bearing health condition.

Processing of lengthy 1D data necessitates an enormous computational time. To handle these issues, an adjustable skimming window mechanism is proposed [102]. This technique aims to achieve (a) efficient training of the network by cutting down repetitive AE signals into small segments, which will eventually yield a good amount of source data for training, (b) handling the issue of fitting the lengthy amount of data to the network as input through segmentation, and (c) stacking segmented AE signals together. This can be useful to further process and generate some identical patterns via imagining, which can be fed to the ADCNN network for classification. If the total length of the AE signal is A_{Lr} , then the total number of the samples A_{Nr} is as follows.

$$A_{Ni} = \left(\frac{A_{Li} - A_{Lf}}{A_{Ls}} \right) + 1 \quad (4.2)$$

Here, A_{Lf} denotes the length of a single frame, f ($f = f_a, f_b, \dots, f_n$) and the step size is A_{Ls} . The complete process of the adjustable skipping window technique is illustrated in Figure 4.4.

In this study, Fast Fourier Transformation (FFT) is applied to each segmented signal. If S_N is the number of total samples, then FFT will perform $S_N \times \log(S_N)$ operations. The reasons behind using this FFT are gathering better analysis of the denoised signals based on the frequency domain

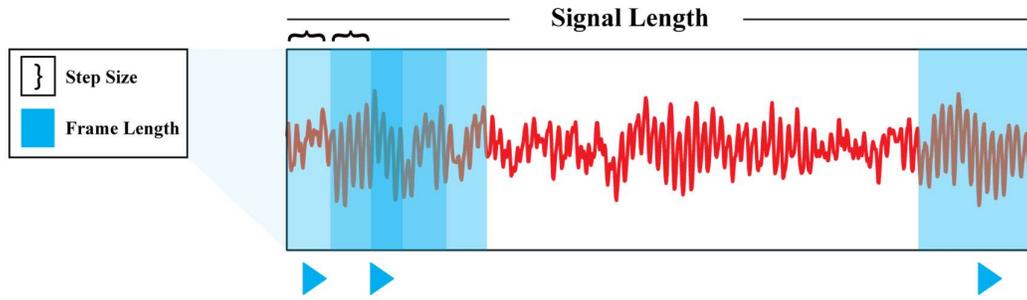


Figure 4.4: Adjustable sliding window technique for the proposed acoustic spectral imaging (ASI).

and creating identical images from stacked signals. For each health type, conversion of denoised signals into frequency domain results in identical visual patterns. In our experiment, it was demonstrated that frequency resolution could provide identical patterns. FFT of each sample is based on the Nyquist theorem, and half of the data points in the frequency spectrum are considered.

Then, the stack of the preprocessed signals forms indistinguishable patterns for each health type. The images are larger when stacking many segmented signals. To tackle this issue, in this study, the height of the stacked signals for generating identical images was fixed based on experiments. If the total number of segmented signals is Z and each portion with a length of W , then the final size of the ASI is $Z \times W$ (*Height* \times *Width*). The produced ASI with the original height might complicate the training process while these ASIs are fed to ADCNN architecture since ASI dimension is still high. Therefore, to create a proper size of ASI images, a number of samples, z , is considered from the Z segmented samples (where $z \in Z, z < Z$ and $Z > 80$) and then $z \times W$ sized images are bunched together to be fed to the network. In this study, $z = 80$ can create a better outcome for ASI from the denoised frequency resolution.

4.2.3 Proposed Fault Diagnosis System with an ADCNN

1) Deep CNN (DCNN) Architecture Overview

CNNs are biologically inspired variants of the multi-layer perceptron that have proved efficient in areas such as image recognition and classification [103, 104]. CNN is generally designed to take advantage of the 2D structure of input [103, 105]. However, the original bearing fault signal is a 1D time-domain signal, from which it is difficult to observe any pattern regarding the bearing defect. Therefore, this study applies an appropriate transformation technique for attaining bearing health states in a 2D visualization. Once a 2D image of the raw signal using ASI analysis is obtained, an ADCNN used for multi-fault classification under variable RPMs. A typical ADCNN for fault diagnosis is shown in Figure 4.5, which is inspired by LetNet-5, in [100]. This proposed ADCNN architecture with an adaptive learning rate is composed of 3 convolutional layers, 3 pooling layers, and 1 fully connected layer, excluding the input, each of which contains trainable parameters (weights).

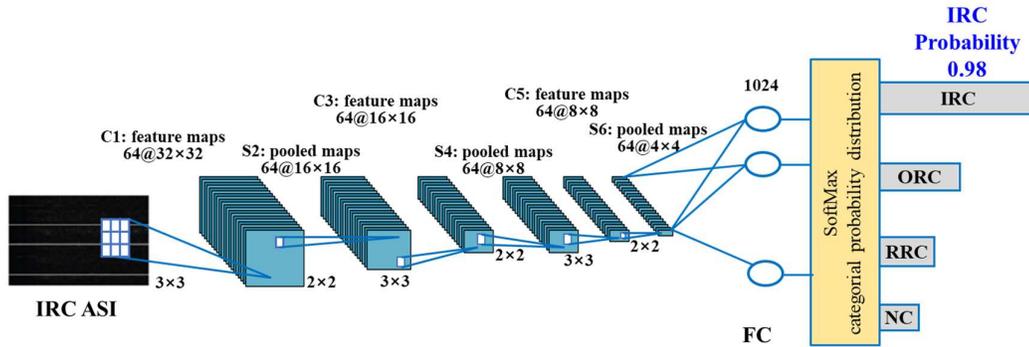


Figure 4.5: The proposed ADCNN architecture used to classify different fault types, where an ASI containing the bearing health state feeds the deep neural network as the input data.

In Figure 4.5, layers CX are the convolutional layers that extract features through convolutional operations, SX are the sub-sampling layers that perform the max-pooling operations, and FX is the fully connected layer, which connects the network to the output layer and is used for classification, where X is the layer index. First, the input image of the ASI is resized to 32×32 pixels to enter the first convolutional layer, $C1$. The different variables and dimensionalities of the networks are listed in Table 4.2.

2) An ADCNN with Back-propagation

In this chapter, the back-propagation (BP) algorithm is applied to update all the trainable parameters on all layers in the proposed ADCNN [103]. In the proposed architecture, the

Table 4.2: Dimensions of the variables used in the proposed convolutional neural network

Layers	Parameters	Remarks	Height	Width	Depth
Input	Reshape <i>ASI</i>		32	32	1
C1	Kernel Size	Filter	3	3	
	Depth	Filter number			64
	Output		32	32	64
S2	Stride		2	2	
	Pooling	Max			
	Output		16	16	64
C3	Kernel Size	Filter	3	3	
	Depth	Filter number			64
	Output		16	16	64
S4	Stride		2	2	
	Pooling	Max			
	Output		8	8	64
C5	Kernel Size	Filter	3	3	
	Depth	Filter number			64
	Output		16	16	64
S6	Stride		2	2	
	Pooling	Max			
	Output		4	4	64
FC	Flattened S6 Output	One-dimensional feature maps	1024		
SoftMax	Target Nodes	OC, IC, RC, and HB	4		

convolutional layers are placed subsequently with sub-sampling layers to progressively build up spatial invariance and to reduce the computational complexity. However, fully connected layers are set in the last step of the architecture to enable consistency with the generic structure, as can be seen in Figure 4.5. In a convolutional layer, the nearby receptive fields of the preceding layer's feature maps are convolved with learnable kernels, and the result is fed into an activation function to generate the output feature map. Each output feature map can be created by combining the convolutional implementations of multiple input maps. Let layer l be the current convolutional layer in a network. In general, the output feature map in that layer is calculated as the following:

$$y_j^l = f\left(\sum_{i \in M_j} y_i^{l-1} * k_{ij}^l + b_j^l\right), \quad (4.3)$$

where j is the j^{th} output feature map in a convolutional layer, i is the i^{th} input feature map, M_j defines a selection of input feature maps in the layer $l-1$, k_j represents a convolutional kernel of feature the j^{th} map, and f is an activation function. The variable b is the bias added to each output map. With different output feature maps, the input feature maps are convolved with different kernels k .

It is important to note that sub-sampling layers help decrease the resolution of the input maps via nonlinear sub-sampling. This study applies maximum pooling in the sub-sampling layer that divides the input image into a set of non-overlapping rectangles and outputs the maximum for each such sub-region. The pooling layer aids in the gradual reduction of the number of parameters to minimize the spatial size of the representation and the computational burden in the network. Hence, the pooling layer also acts to control overfitting. Therefore, this study periodically inserts a pooling layer between successive convolutional layers, as can be seen in Figure 4.5. Furthermore, the pooling operation provides another useful form of translational invariance. The pooling layer is defined as follows:

$$y_j^l = f(\varphi_j^l SS(y_j^{l-1}) + b_j^l), \quad (4.4)$$

where $SS(\cdot)$ defines a sub-sampling function. Generally, this function computes the maximum of each distinct 2-by-2 block in the input feature map, such that the output feature map is 2 times smaller than the number of rows and columns in the input feature map. Each output map is given its trainable coefficient (φ) and trainable bias (b).

The final step of the proposed ADCNN architecture is the design of the fully connected (FC) layer: the last few layers (closest to the outputs) are a fully connected 1-D layer. The output of this layer is

$$o^l = f(y^l w^l + b^l), \quad (4.5)$$

where the output activation function $f(\cdot)$ is the logistic function and w consists of trainable weights.

As the BP algorithm is applied to train the ADCNN, the gradients of the loss function for all trainable weights in all layers are calculated during backward operation of the BP. However, it is significant to define an appropriate objective function. Thus, a squared-error loss function is applied in this chapter to address the objective function. Eq. (4.5) represents the loss function after training a single sample, such as i . The overall loss function is defined by summing the loss functions of the individual samples, as below:

$$\text{cost}(w) = \frac{1}{2} \sum_{i=1}^m (o_i^l - t_i^l)^2, \quad (4.6)$$

where t_i^l defines the target output value of the i^{th} pattern.

Suppose a point w is used to find the next weight point ($w+1$) to find a minimizer, and the step starts from w and moves in increments of $\alpha \frac{\partial}{\partial w} \text{cost}(w)$, as in Eq. (16), where α is a positive scalar step size.

$$w := w - \alpha \frac{\partial}{\partial w} \text{cost}(w) \quad (4.7)$$

That weight update process in Eq. (4.7) is called a stochastic gradient descent (SGD) algorithm [103]. It is assumed that the gradient in the SGD will vary as the search continues and tends toward zero as it approaches the minimizer. As such, the step sizes can be either small or large. This first approach (*e.g.*, a small step size) results in a computationally complex path to reach a minimizer; however, a second approach (*e.g.*, a large step size) could result in a more zigzag-like path to the minimizer. This approach is computationally inexpensive and simple but could become trapped in a local minimizer; these points taken together indicate that there is no globally optimal value of the step size.

To evaluate the above problem in the SGD, this study applies an adaptive moment estimation that combines the advantages of the adaptive gradient (AdaGrad)— which works well with sparse gradients — and a root-mean-square propagation (RMSProp) — which works well in nonstationary settings [106]. The main idea is to maintain exponential moving averages of the gradient and its square in each update, proportional to the average gradient and its square as follows:

$$w := w - \alpha \frac{M_t}{\sqrt{R_t + \epsilon}} \quad (4.8)$$

where,

$$M_t = \beta_1 M_{t-1} + (1 - \beta_1) \frac{\partial}{\partial w} \text{cost}(w) \quad (4.9)$$

$$R_t = \beta_2 R_{t-1} + (1 - \beta_2) \frac{\partial^2}{\partial w^2} \text{cost}(w) \quad (4.10)$$

where M_t is the 1st-moment bias correction, R_t is the 2nd-moment bias correction, and the decay rates are small (i.e., β_1 and β_2 are close to 1). The weight update process in Eq. (4.8) provides a method to signify a nearly optimal learning rate selection.

Table 4.3: Details of the considered working conditions with the same health types

Datasets	Health Type	Shaft Speed (rpm)	Number of Singles, with Sampling Frequency (250K Hz) and Signal Length (1 sec)	Crack Size Length (mm)		
Dataset 1	Normal Condition (NC)	250	110	6		
	Inner Race Fault Condition (IRC)	250		6		
	Outer Race Fault Condition (ORC)	250		6		
	Roller Raceway Fault Condition (RRC)	250		6		
Dataset 2	Normal Condition (NC)	300		110	6	
	Inner Race Fault Condition (IRC)	300			6	
	Outer Race Fault Condition (ORC)	300			6	
	Roller Raceway Fault Condition (RRC)	300			6	
Dataset 3	Normal Condition (NC)	350			110	6
	Inner Race Fault Condition (IRC)	350				6
	Outer Race Fault Condition (ORC)	350				6
	Roller Raceway Fault Condition (RRC)	350				6

4.3 Results and Discussion

4.3.1 Dataset description

For an investigation of the appropriateness of the proposed fault diagnosis, (ASI + ADCNN), scheme including ASI-based RPM invariant bearing health state visualization and

ADCNN based RPM invariant bearing fault classification, extensive simulation, and experimental analyses are presented in this section. The benchmark AE dataset of bearing faults was used to conduct the experiments. This chapter used three different RPMs (250, 300, and 350) and recorded 110 signals for each health type (e.g., NC, IRC, ORC, and RRC) at each RPM. The details of the dataset are given in Table 3.

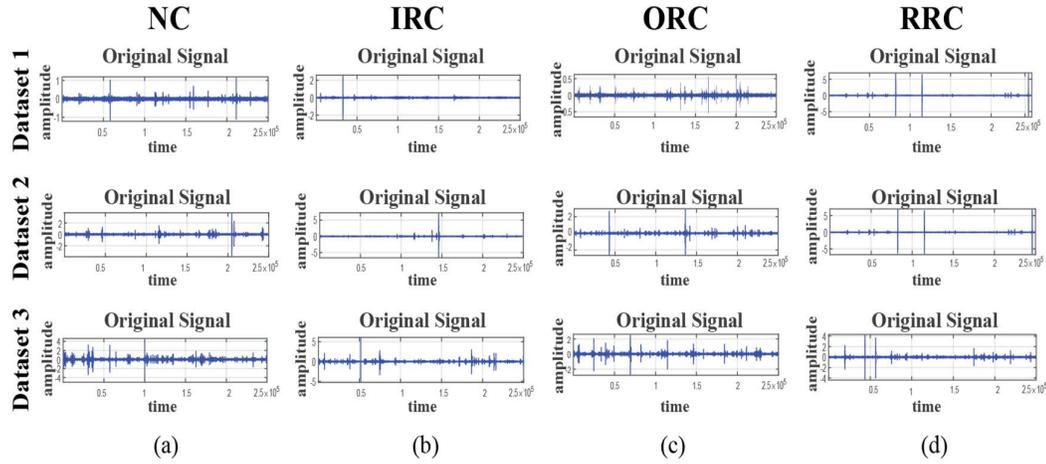


Figure 4.6: Different raw signals for different health conditions: (a) NC, (b) IRC, (c) ORC, and (d) RRC

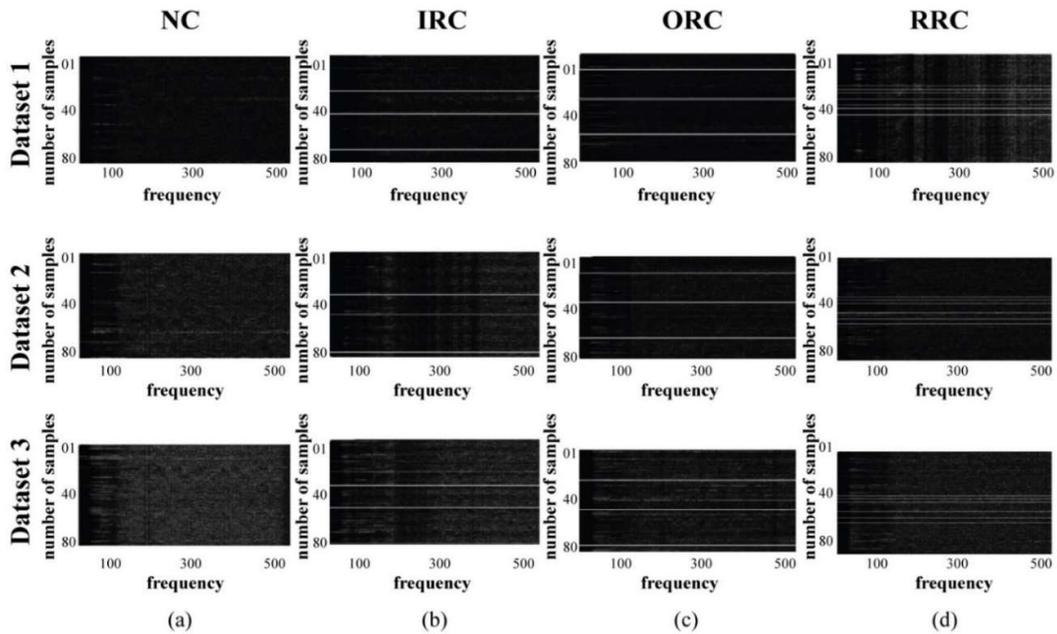


Figure 4.7: Different ASI for different health conditions: (a) NC, (b) IRC, (c) ORC, and (d) RRC

4.3.2 Performance Analysis of Acoustic Spectrum Imaging (ASI)

As explained in Section 4.1, the 1D signal reveals no discernible patterns regarding faults and almost no information about RPM variance, as can be seen in Figure 4.6. Therefore, 2D-based ASI to represent the bearing health state is developed. To visualize the effectiveness of ASI, this chapter utilizes three datasets with four fault types (NC, IRC, ORC, and RRC) at various operating conditions (see Table 4.3).

Figure 4.7 presents the results of the 2D-based ASI for bearing health state representation. According to the results in Figure 4.7, it is apparent that the frequency band as a function of time for different health conditions yields different patterns. In the considered dataset, though the rpm varies, the pattern remains identical and distinguishable for different health types. An interesting point to note that it can be seen there is no defect frequency and its harmonics in the case of healthy bearing (e.g., NC) while other bearing conditions (e.g., IRC, ORC, and RRC) clearly show defect frequencies and their harmonics depending on bearing fault conditions as shown in Figure 4.7, which also validates the effectiveness of ASI. In short, the same type of health conditions generates the same kind of patterns where they remain identical to the others. This is because the 1D AE signal is proposed in a way so that all time information is stacked together for converting into a 2D image to utilize the distribution of energies in various frequency bands.

4.3.3 Diagnostic performance of the proposed method

As proposed ASI is highly effective to visualize the performance of bearing health representation, to further take full advantage of the deep learning architecture, ASIs are used as an input to the ADCNN for diagnosing bearing faults under variable speeds.

To authenticate the performance of the proposed ADCNN method, it is essential to divide the dataset into appropriate training and testing for preparing RPM variant scenarios. In the first scenario, dataset 1 is utilized to train the network for storing the knowledge, whereas datasets 2 and 3 utilize the stored knowledge to complete the classification test. From datasets 2 and 3, 20% of the data is utilized for adapting the network to use the earlier gathered knowledge. In the second scenario, dataset 2 is used for gathering the knowledge and datasets 3 and 4 are used for testing. Similarly, in scenario 3, dataset 3 is used for knowledge gathering, whereas datasets 1 and 2 are used for classification. In each scenario, one dataset is known to the network and the other datasets utilize the learning knowledge to make the learning faster and efficient by maintaining the TL principle to archive the accuracy. For a reliable comparison of diagnostic performance, this study measures the sensitivity (SEN) or true positive rate (TRP), and average classification accuracy

(ACA), and overall classification accuracy (OCA), as in [42]. SEN is the percentage of samples of a particular class that are correctly classified and is calculated as follows.

$$SEN = \frac{T_{TP}}{T_{TP} + T_{FN}} \times 100 (\%). \quad (4.10)$$

Where T_{TP} is the number of samples from a particular class that is correctly classified. T_{FN} is the numbers of samples from a particular class that are not correctly classified. The final result is calculated in percentage to determine the final class-wise result. After calculating the final SEN of each health type, this study measured the average classification accuracy (ACA) for a particular target dataset by following Eq. (4.11).

$$ACA = \frac{\sum SEN}{\sum T_c} \quad (4.11)$$

where $\sum SEN$ is the summation of class-wise accuracy for a particular target dataset.

Finally, the scenario-based overall classification accuracy (OCA) is measured to get the overview of the final experimental analysis, which is defined as underneath,

$$OCA = \frac{\sum SD_{ACA}}{\sum TS_{Scenario}} \quad (4.12)$$

Here, $\sum SD_{ACA}$ defines the summation of ACA of source dataset, and $\sum TS_{Scenario}$ defines the total number of source dataset present in an individual scenario. Details classification results are presented in Table 4.4. To explain the speed (RPMs) variation situation, this chapter prepared three scenarios to measure the classification performance. For example, in scenario 1, target dataset is dataset 1 and rest two datasets such as dataset 2 and 3 are source dataset. According to result in Table 4.4, sensitivity (SEN) of each fault type is 98.30%, 84.23%, 95.77%, and 92.87% for NC, IRC, ORC, and RRC respectively for source dataset 2 in the scenario 1. And ACA score is 92.79%. Likewise, this chapter calculates the classification accuracy of dataset 3, and this chapter attains 95.20% of ACA score for dataset 3. Finally, the OCA score is 93.995% for scenario 1, which

is calculated through Eq. (4.12). In addition, it can be seen that the average accuracy of the different health types for target conditions are 95.73%, 87.14%, 95.04% and 92.78% for NC, IRC, ORC, and RRC, respectively. The average accuracy of OCA is 92.67%.

For each scenario, the test dataset is five times higher than the training set which means that ADCCN can work effectively even when the operating conditions change. To fine-tune the network and to obtain classification accuracy, 300 fine-tuned epochs are used.

In addition, to prove the robustness of our proposed method (ASI + ADCNN), the proposed method is compared with three state-of-the-art approaches for bearing fault diagnosis, i.e., AE based classical approach [107], a deep transfer learning method based on neural networks with raw 1D signal [102], and a deep learning technique with 2D energy spectrum map as input [108]. To provide

Table 4.4. Diagnostic performance of the proposed model under different scenarios

Scenario	Target Dataset	Source Dataset	Sensitivity (SEN) (%)				Average Classification Accuracy - ACA (%)	Overall Classification Accuracy – OCA (%)
			NC	IRC	ORC	RRC		
1	Dataset 1	Dataset 2	92.30	82.23	93.77	92.87	90.29	91.74
		Dataset 3	97.81	87.37	94.28	93.34	93.20	
2	Dataset 2	Dataset 3	97.38	86.85	94.84	93.53	93.14	93.62
		Dataset 1	95.68	88.27	96.87	95.58	94.10	
3	Dataset 3	Dataset 1	96.24	85.99	97.64	94.23	93.52	92.32
		Dataset 2	95.07 7	85.01	93.85	90.59	91.13	
Average Accuracy			95.74	85.95	95.20	93.35	92.45	

a fair comparison, this chapter have tested our proposed method and the conventional approaches using the same dataset. The proposed method that combines acoustic spectral imaging (ASI)— a new 2D imaging technique of 1D raw signal— with a deep learning technique based on convolution neural network (CNN) for diagnosis bearing faults under variable speeds conditions. In [107], first Hilbert-Hung Transform (HHT) is applied to identify an informative frequency-band and then salient features are extracted in the selected frequency band. Finally, a classical machine learning approach such as a support vector machine (SVM) is combined with salient features for calculating classification accuracy. In [108], 1D raw signals are transformed into 2D spectral energy maps (SEMs) and these SEMs are used as input to convolution neural network (CNN) to learn the optimal

features automatically for diagnosing bearing health conditions under variable operating speeds. In [102], raw 1D signals are fed to the transfer learning (TL)-based CNN to extract the features automatically. Table 4.5 compares the results of the proposed scheme with than that of other approaches. According to the result in Table 5, it can be clearly seen that the proposed method (ASI + ADCNN) outperforms its counterparts. In the AE based classical approach [46], the selected features contain some overlapping points, so that the performance of the classifier degrades. Similarly, in [40] and [42], much redundant information and irrelevant feature overlap reduce the overall classification performance. The proposed approach outperforms the state-of-the-art, yielding an overall performance improvement of 16.55%, 1.15%, and 20.21% in all scenarios against [107], [102] and [108] respectively.

Table 4.5: Comparison of the classification accuracy

Scenario	Method	Sensitivity - SEN (%)				Average Classification Accuracy ACA (%)	Improvement (%)
		NC	IRC	ORC	RRC		
1	[107]	77.25	65.3	76.73	74.47	73.44	16.55
	[102]	94.23	84.59	91.7	92.83	91.84	1.15
	[108]	71.21	62.2	67.12	70.58	70.78	20.21
	Proposed	99.06	86.8	96.03	94.11	93.99	-
2	[107]	78.11	69.32	77.6	75.42	75.11	20.51
	[102]	92.44	85.9	94.82	93.2	91.59	4.03
	[108]	72.56	63.62	70.59	74.22	70.23	25.39
	Proposed	98.53	89.56	97.86	96.56	95.62	-
3	[107]	79.69	67.7	77.92	76.41	75.43	18.97
	[102]	94.5	86.95	93.45	90.3	91.3	3.1
	[108]	72.44	66.31	70.23	73.39	70.59	23.81
	Proposed	98.61	88.06	97.25	93.68	94.4	-

In addition to average classification accuracy (avg.) and sensitivity-based performance measurement, this chapter also calculated precision and F1 score to verify the reliability of the proposed method. Precision or positive predictive value (PPV is calculated in the following equation,

$$PPV = \frac{TP}{TP + FN} \tag{4.13}$$

Where, TP is true positive counts and FN is false negative counts. F1 score or the harmonic mean of precision and sensitivity is calculated as follows,

$$F1 = 2 \cdot \frac{PPV \cdot TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN} \quad (4.14)$$

Table 4.6 compares the results of PPV and F1 of the proposed method and three state-of-the-arts in different scenarios. In the results, it can be seen the precision and F1 of the proposed method are also significantly high, which ensures the reliability and robustness of the proposed method.

Table 4.6: Comparison of the classification accuracy in various performance matrices

Scenario	Method	Fault Type wise SEN/PPV Values (%)				ACA/F1 Scores (%)
		NC	IRC	ORC	RRC	
1	[107]	77.25/82.0	65.3/73.5	76.73/75.50	74.47/81.50	73.44/78.12
	[102]	94.23/91.30	84.59/89.35	91.7/90.50	92.83/92.50	91.84/90.91
	[108]	71.21/70.60	62.2/70.50	67.12/68.67	70.58/75.60	70.78/71.6
	Proposed	99.06/98.97	86.8/91.5	96.03/95.05	94.11/95.53	93.99/95.26
2	[107]	78.11/80.10	69.32/72.60	77.6/72.85	75.42/74.50	75.11/75.01
	[102]	92.44/89.50	85.9/90.20	94.82/91.50	93.2/93.50	91.59/91.17
	[108]	72.56/75.50	63.62/64.32	70.59/73.60	74.22/78.60	70.23/75.0
	Proposed	98.53/99.0	89.56/93.40	97.86/96.45	96.56/97.31	95.62/96.23
3	[107]	79.69/80.0	67.7/75.40	77.92/78.50	76.41/73.60	75.43/76.87
	[102]	94.5/94.0	86.95/90.0	93.45/91.80	90.3/90.0	91.3/91.45
	[108]	72.44/75.0	66.31/65.50	70.23/75.6	73.39/80.0	70.59/73.9
	Proposed	98.61/100	88.06/87.0	97.25/96.69	93.68/94.05	94.4/94.43

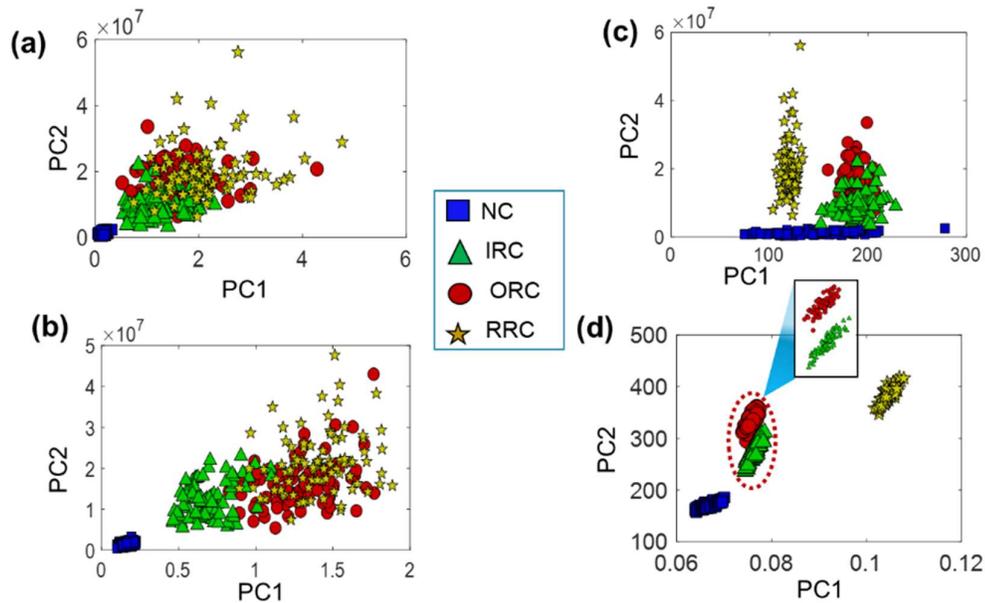


Figure 4.8: 3-dimensional visualization of different features using KPCA: (a) [108], (b) [107], (c) [102], and (d) Proposed.

In order to demonstrate the superiority of the proposed method in feature learning, Figure 4.8 displays two-dimensional visualization of features learned by different methods. As is shown in Figure. 4.8 (a) and (b), the projection of data is mostly overlapped among different fault classes. modes. The features of each fault modes learned by the proposed method compactly cluster together, which are much more distinguishable and identifiable than other methods.

4.4 Conclusions

This study presents two-dimensional acoustic frequency spectral imaging (ASI) with a transfer learning (ADCNN)-based fault diagnosis method that adds a new dimension to bearing fault diagnosis, which is invariant to both random and deliberate differences of the shaft speed. In traditional approaches, the diagnosis of the bearing is mainly based on detecting defect frequencies. These techniques have certain challenges, for example, a non-stationary shaft speed creates an impact on defect frequencies, and variations of the shaft speeds bring significant and tiny variations on defect frequencies. In addition to this, the conventional feed-forward neural network mechanisms lack autonomous feature extraction for improving classification performance as in the CNN and DL. This study validated our proposed method (ASI + ADCNN) by using the health images of four different health conditions for three different rotational speeds. The proposed fault diagnosis architecture yields an average accuracy of 92.45%, which establishes this proposed method as invariant to variations of shaft speed. In addition, the proposed framework outperforms three the state-of-the-arts, yielding an overall performance improvement of 16.55%, 1.15%, and 20.21% in all scenarios respectively.

Chapter 5

Data-Driven Prognostic Scheme for Rolling-Element Bearings Using a New Health Index and Variants of Least-Square Support Vector Machines

5.1 Introduction

The desire for correct fault diagnosis and accurate predictive prognosis has been around since the industrial revolution and would benefit complex machinery such as aircraft, spacecraft, automobiles, wind turbines, and pumps by providing an early indication of impending failure and forecasting remaining useful life (RUL) [3, 18, 109]. Rolling-element bearings (REBs) are crucial components in these machines and are frequently prone to failure [110-112]. Therefore, prognostics and system health management (PHM) of REBs is vital to maximize component uptime and minimize operation and maintenance (O&M) costs. Intelligent prognosis capability alludes specifically to the phases involved with onset anomaly detection in the acquired data, as the time-to-start (TTS) point for RUL estimation, predicting future degradation trends of faults, and estimating the RUL. In the PHM research area, fault prognosis can be performed either in a model-based or data-driven manner.

When using model-based techniques, it is important to understand true dynamic models of systems for computing the time-dependent damage rates and the process of their accumulation.

Model-based techniques include autoregressive moving-average (ARMA) approaches [113], physics-based models [114], Kalman filters [13, 115], particle filters, and empirical model-based methods. In [115], the fault degradation trend is assumed to follow an exponential model for RUL estimation, and unknown model parameters are calculated from the measured data through the expected maximization (EM) algorithm. Saidi et al. [12] employed a physics-based health index (HI) integrated with the Paris model for estimating the RUL of REBs. Despite their usefulness, model-based techniques are best suited when an accurate mathematical model of the system can be constructed. In real-world applications, these processes are too multifaceted, and constructing an inclusive model and accounting for all coupling effects is complex and practically impossible. Therefore, hard-coded models are mainly system specific and cannot be generalized for diverse systems.

Alternatives to model-based prognostic techniques are data-driven approaches (i.e., data mining or machine learning approaches), which directly learn from the routinely monitored machine operating data (e.g., vibration data) [116, 117]. Neuro-fuzzy models [118, 119], the hidden Markov model (HMM) [120], the Gaussian mixture model (GMM), and linear/polynomial regression models have been substantially employed as data-driven methods for RUL prediction in various mechanical and electrical machines. In [121], the authors applied an artificial neural network (ANN) to learn about impending bearing failure to predict the RUL. Recently, Ahmad et al. [110] proposed the use of polynomial regression for RUL estimation based on a root mean square (RMS)-based bearing health index. Even though these data-driven approaches circumvent many of the limitations facing model-based techniques, their efficiency is strongly dependent on both the quantity and the quality of system operation data. Furthermore, these schemes are black-box processes that ignore the underlying physics of the overall process; it is also impossible to incorporate any available prior information. Alternatively, the development of support vector machines (SVMs) deals with the sound theory first, then implementation and experimentation. Recently, least-square SVM (LSSVM) [122], which applies equality instead of inequality constraints and can be easily extended to the probabilistic interpretation of its outputs as well as to a recurrent regression model, has been widely applied for problems of fault diagnosis and prognosis. Thus, this chapter modifies LSSVM as Bayesian inference-based one-class LSSVM (Bayesian-OCLSSVM) for accurate TTS point detection and as a recurrent least-square support vector regression (Recurrent-LSSVR) model for robust RUL estimation by predicting future HI values.

In the data-driven PHM schemes for REBs, there are still several important issues that must be addressed, such as constructing an appropriate health index (HI) for a bearing that demonstrates a smooth degradation trend with progression of the mechanical failure, determining the anomalies

in the HI trend accurately to define the TTS point, and defining a regression model capable of predicting future values of the HI for accurate RUL estimation at any point. To the best of our knowledge, most of the literature related to bearing PHM using data-driven techniques is mainly focused on HI construction, anomaly detection, or regression modeling; however, no previous studies address all three issues simultaneously, despite the fact that these are all integral parts for estimating a robust RUL. Therefore, this chapter presents a new HI by defining a degree-of-defectiveness (DD) metric in the frequency-domain of bearing time-domain signal (referred to as DD-HI) with a robust PHM module consisting of Bayesian-OCLSSVM for detecting abnormal instances in DD-HI sequences, Recurrent-LSSVR for predicting the degradation trend, and RUL estimation.

Properly designing an anomaly detector that is capable of detecting early raising faults in bearings and providing important information about the time (from which the estimation of the RUL can be started) is essential. This is the case because robust detection of TTS affects the accuracy of the RUL estimation procedure. Our anomaly detection method, i.e., Bayesian-OCLSSVM, is based on a probabilistic model that combines Bayesian inference with an OCLSSVM machine learning technique. The original concept of LSSVM was introduced by Suykens [18], and its main advantage over conventional SVM techniques is that the LSSVM model can be obtained by solving a linear programming problem with equality constraints that help us generate the probabilistic decision output with Bayesian inference. This anomaly detection method using Bayesian-OCLSSVM can be trained with a small amount of data, corresponding to the normal operating conditions, in an online fashion.

Once an anomaly is detected, the proposed prognostic framework activates the recurrent-LSSVR module to predict the future trends of DD-HI and RUL calculation. When a bearing starts to degrade, it may follow linear, exponential, high-order polynomial, or stochastic trends. Describing the degradation process using a fixed model, such as a linear function, exponential function [115], or polynomial function [110], may compromise the overall accuracy of RUL estimation. For example, a regression model with a high-degree polynomial function may fit non-linear training data much better than a linear function. However, this model can severely overfit the data (e.g., 100-degree polynomial function). At the same time, linear regression is still prone to underfitting the data. Consequently, it is essential to explore the effect of the polynomial order concerning the goodness-of-the-fit to find a suitable model. Therefore, recurrent-LSSVR is a completely data-driven technique that applies a recurrent weighted LSSVR model to predict future data trends.

In addition to the proposed new DD-HI and developed PHM solution containing Bayesian-

OCLSSVM-based anomaly detector and Recurrent-LSSVR-based regression model, this chapter also introduces a dynamic thresholding technique, that calculates the gradient of a linear fit model in the updated window, to determine the time of bearings expected failure.

The layout of this chapter is outlined as follows. In Section 5.2, this chapter discusses the proposed prognostic framework, including DD-HI construction, Bayesian-OCLSSVM, and Recurrent-LSSVR. Section 5.3 introduces the experimental setup and data-acquisition system. The experimental results and relevant discussions are provided in Section 5.4. Finally, Section 5.5 provides our concluding remarks.

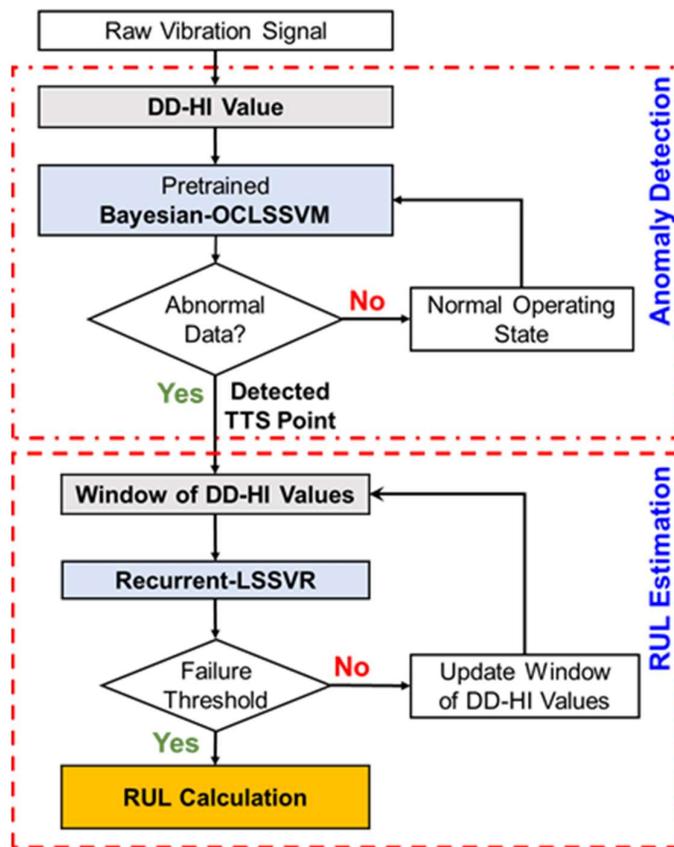


Figure 5.1: Proposed prognosis framework for estimating the remaining useful life (RUL).

5.2 Proposed Methodology

Figure 5.1 illustrates the integrated architecture of the proposed PHM solution including a new health index (e.g., DD-HI) and variants of least-square support vector machines (LSSVM).

5.2.1 Degree-of-Defectiveness-based HI (DD-HI)

Very few statistical feature parameters (e.g., the RMS, mean value, variance, kurtosis, and skewness of the time-domain signal) have been investigated as bearing HIs [18] with various forms of adjustments (e.g., smoothing, rectification, and filtering). This kind of subjective adjustment lacks the actual degree of defects in the HI. Apart from time-domain features, frequency-domain properties have been significantly investigated for developing bearing degradation trends over the last few decades [123, 124]. A proven frequency-domain analysis is the envelope power spectrum (EPS), which is capable of detecting localized defects in bearings [3]. Whenever bearing failure appears, there are specific defect frequencies at

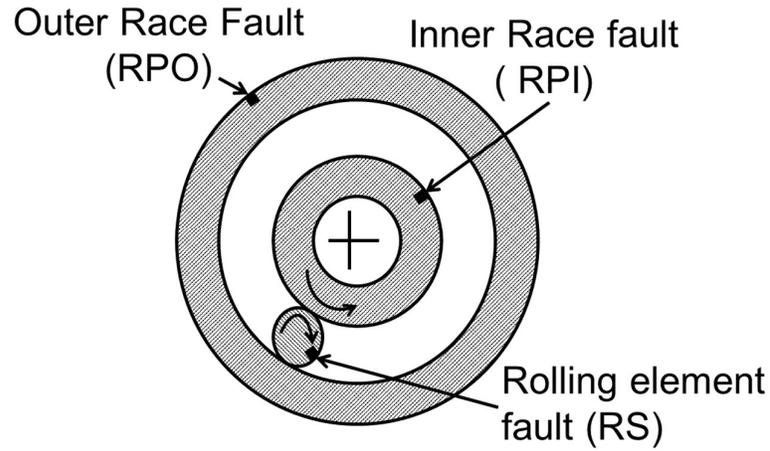


Figure 5.2: Bearing configuration with possible fault locations.

which faulty symptoms are highly detectable. These faulty symptoms can be detected at one of the defect frequencies listed in Eqs. (5.1) -(5.4), depending on how the roller strikes a point of local failure [3]:

$$RPO = \frac{nf_r}{2} \left(1 - \frac{R_d}{P_d} \cos \theta \right), \quad (5.1)$$

$$RPI = \frac{nf_r}{2} \left(1 + \frac{R_d}{P_d} \cos \theta \right), \quad (5.2)$$

$$RS = \frac{P_d f_r}{2R_d} \left[1 - \left(\frac{R_d}{P_d} \cos \theta \right)^2 \right], \quad (5.3)$$

$$FTF = \frac{f_r}{2} \left[1 - \left(\frac{R_d}{P_d} \cos \theta \right) \right]. \quad (5.4)$$

Here, RPO is the roller pass frequency for the outer race fault, RPI is the roller pass frequency for the inner race fault, RS is the roller spin frequency, and FTF is the fundamental train frequency for the fault. These frequencies depend on the shaft speed (f_r), the number of rolling elements (n), the contact angle from the radial plane (θ), the roller diameter (R_d), and the pitch diameter (P_d). The intuitive meanings of these defect frequencies can be understood from Figure 5.3.

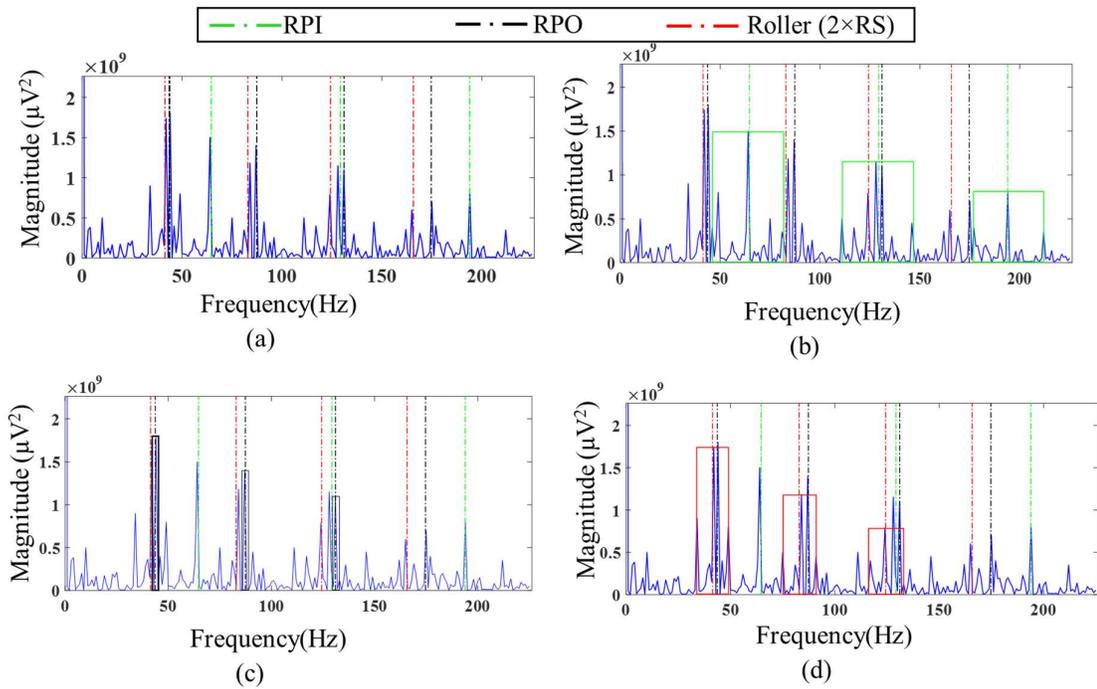


Figure 5.3: (a) An envelope power spectrum of a combined fault, and an illustration of the defect range calculation up to three harmonics (using equations (8), (9), and (10)) in the envelope power spectrum on which RMS features are calculated for (b) outer, (c) inner, and (d) roller defects, respectively.

The EPS is highly effective in detecting bearing defect frequencies, such as RPO, RPI, and RS, depending on their localized defects. Therefore, this chapter utilizes these defect frequencies (up to their third harmonics) to compute the proposed health index, DD-HI. This process includes two

steps. First, narrow-band degree-of-defectiveness (DD) regions are defined around the RPO, RPI, and RS frequency bins by analyzing the dynamics of bearing failure. Then, RMS values are calculated in the DD region for health index construction; this is referred to as DD-HI. To explain the dynamics of the bearing, the radial load significantly influences the force of the impact yielded by rolling over a defect. The outer race is a stationary part in the bearing; thus, a defect on the outer race is exposed to the same force whenever a rolling element passes the location of this defect. However, failure on the inner race has a varying force since it rotates at almost the shaft speed. As a result, all the harmonics of the RPI are amplitude-modulated by the shaft speed (i.e., f_r). Likewise, the value of $2 \times RS$ produced by a roller element defect is amplitude-modulated by the RPM of the cage (i.e., FTF). Theoretically, amplitude modulation, which appears due to the inner- or roller-related bearing defects, creates sidebands that are spaced apart by the modulation frequency (i.e., f_r or FTF) and centered around the RPI or $2 \times RS$. Moreover, it is common to observe random variations (RVs) in the calculated bearing defect frequencies on the order of 1–2% [3]. To define narrow-band DD windows around the three defect frequencies (i.e., RPO, RPI, and RS), the following formulations are defined:

$$(1 - RV) \cdot (RPO_h) \text{ to } (1 + RV) \cdot (RPO_h), \text{ from}$$

$$(1 - RV) \cdot (RPI_h - 2 \cdot f_r) \text{ to } (1 + RV) \cdot (RPI_h + 2 \cdot f_r), \text{ and} \quad (5.5)$$

$$\text{from } (1 - RV) \cdot (2 \times RS_h - 2 \cdot FTF) \text{ to } (1 + RV) \cdot (2 \times RS_h + 2 \cdot FTF).$$

Here, RV is 2% of each of these defect frequencies, to properly fit the windows, and h is the order number of the defect frequency harmonic ($h=1, 2,$ and 3 in this chapter). Then, this chapter computes the RMS values, such as RMS_{RPOh} , RMS_{RPIh} , and $RMS_{2 \times RS_h}$, for RPO, RPI, and RS narrow-band DD regions, respectively. Our proposed DD-HI is the arithmetic average of these RMS values, as shown below:

$$DD - HI = \frac{1}{3} \sum_h (RMS_{RPOh} + RMS_{RPIh} + RMS_{2 \times RS_h}). \quad (5.6)$$

The health index defined in Eq. (5.6) is more efficient than existing time-domain statistical parameters since this process considers the actual dynamics and physics of failures (POF) appearing in REBs. Thus, it is expected that the proposed DD-HI can forecast a better result when estimating the RUL of bearings.

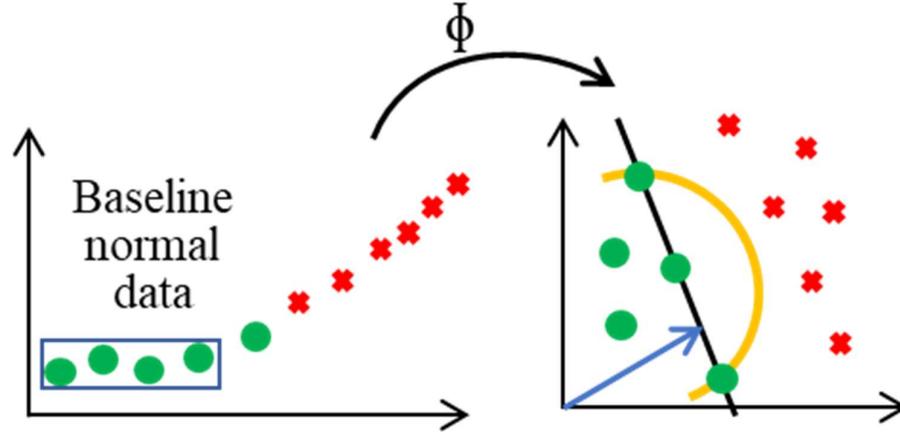


Figure 5.4: One-class, least-square support vector machine (OCLSSVM).

5.2.2 Bayesian-OCLSSVM for Anomaly Detection

This section introduces a Bayesian-OCLSSVM framework and explains how it is utilized for detecting anomalies in bearing data sequences. In classical SVM, the objective is to find the optimal hyperplane that can separate the target class with a maximum margin, which is expressed as follows:

$$\begin{aligned}
 f(x) &: \langle w, \Phi(x) \rangle + b \\
 \text{s.t. } & f(x) > 0,
 \end{aligned}
 \tag{5.7}$$

where w is the weight vector, b is the bias term, and $\Phi(\cdot)$ is a high-dimensional mapping function that is usually replaced by a kernel function. The objective function of classical SVM in Eq. (5.7) solves an inequality constraint but lacks probabilistic interpretation. Instead, this chapter utilizes OCLSSVM, which implies equality constraints and minimizes the optimization function in Eq. (5.8). An illustration of OCLSSVM is given in Figure 5.4.

$$\begin{aligned}
 L(w, \xi, p) &= \frac{1}{2} \|w\|^2 - p + \frac{1}{2} C \sum_{i=1}^n \xi_i^2 \\
 s.t. \langle w, \Phi(x_i) \rangle &= p - \xi_i, \\
 \xi_i &= p - w \cdot \Phi(x_i).
 \end{aligned} \tag{5.8}$$

The OCLSSVM in Eq. (8) aims to define the hyperplane that has the maximal distance $p/\|w\|^2$ from the origin, with respect to which the sum of the squares of errors ξ_i^2 is minimized. To obtain a dual formula, the problem in Eq. (8) can be solved using Lagrangian multipliers, as shown below:

$$L = \frac{\|w\|^2}{2} - p + \frac{C}{2} \sum_{i=1}^n \xi_i^2 - \sum_{i=1}^n \alpha_i (\Phi(x) \cdot w + \xi_i - p). \tag{5.9}$$

Setting the first partial derivatives of Eq. (5.9) to zero with respect to w, ξ_i, p , and α_i , one obtains the following relations:

$$\begin{aligned}
 \frac{\partial L}{\partial w} = 0 &\Rightarrow w = \sum_{i=1}^n \alpha_i \Phi(x_i), \\
 \frac{\partial L}{\partial \xi_i} = 0 &\Rightarrow C \xi_i = \alpha_i, \\
 \frac{\partial L}{\partial p} = 0 &\Rightarrow \sum_{i=1}^n \alpha_i = 1, \\
 \frac{\partial L}{\partial \alpha_i} = 0 &\Rightarrow w^T \Phi(x_i) + \xi_i - p = 0.
 \end{aligned} \tag{5.10}$$

Substituting the first three equations into the last one in (5.10) yields:

$$\sum_{i,j=1}^n \alpha_i \Phi^T(x_i) \cdot \Phi(x_j) + \frac{\alpha_i}{C} - p = 0. \tag{5.11}$$

With the linear constraint for α_i in Eq. (5.10), Eq. (5.11) can be rewritten as a set of linear equations:

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \Phi + \mathbf{I}/C \end{bmatrix} \begin{bmatrix} -p \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix}, \quad (5.12)$$

where \mathbf{a} is the column vector of Lagrangian multipliers $[\alpha_1, \dots, \alpha_n]$, Φ is the Gram matrix with the (i, j)-th entry in $k(x_i, x_j)$, and \mathbf{I} is the identity matrix with compatible lengths for all-one and all-zero column vectors.

Because the solution of Eq. (12) holds a linear equation, to allow for its probabilistic interpretation, this chapter apply Bayesian inference. For a given dataset D , the hyperparameters μ, ξ of Φ and the model parameters w and b can be estimated by maximizing the posterior probability $p(w, b | D, \log \mu, \log \xi, \Phi)$. Applying the Bayesian rule, the posterior probability is obtained as follows:

$$\begin{aligned} & p(w, b | D, \log \mu, \log \xi, \Phi) \\ &= \frac{p(D | w, b, \log \mu, \log \xi, \Phi)}{p(D | \log \mu, \log \xi, \Phi)} p(w, b | \log \mu, \log \xi, \Phi), \end{aligned} \quad (5.13)$$

where $p(D | \log \mu, \log \xi, \Phi)$ is a regulation operator. In addition, it is reasonable to consider that w, b are statistically independent. Thus, one can write $p(w, b | \mu, \xi, \Phi) = p(w | \mu, \xi, \Phi) p(b | \mu, \xi, \Phi)$. Additionally, this can be simplified to $p(w, b | \mu, \xi, \Phi) = p(w | \mu, \Phi) p(b | \sigma_b, \Phi)$, where this study lets $\sigma_b \rightarrow \infty$ to approximate a uniform distribution. When $\sigma_b \rightarrow \infty$, this study gets:

$$\begin{aligned} p(w, b | \log \mu, \Phi) &= \left(\frac{\mu}{2\pi}\right)^{n_h/2} \exp\left(-\mu \frac{1}{2} w^T w\right) \frac{1}{\sqrt{2\pi}\sigma_b} \exp\left(-\frac{1}{2} \frac{b^2}{\sigma_b^2}\right) \\ &\propto \left(\frac{\mu}{2\pi}\right)^{n_h/2} \exp\left(-\mu \frac{1}{2} w^T w\right). \end{aligned}$$

Assuming that the data are independent, the equation for the likelihood can be written as follows:

$$\begin{aligned}
 p(D | w, b, \mu, \xi, \Phi) &= \prod_{k=1}^N p(x_k, y_k | w, b, \xi, \Phi) \\
 &\propto \prod_{k=1}^N p(e_k | w, b, \xi, \Phi) \\
 &= \prod_{k=1}^N \sqrt{\frac{\xi}{2\pi}} \exp\left(-\frac{1}{2} \xi e_k^2\right).
 \end{aligned} \tag{5.14}$$

Then, by combining Eqs. (5.13) and (5.14) and neglecting all constraints, the Bayes rule of inference becomes: $p(w, b | D, \log \mu, \log \xi, \Phi) \propto \exp\left(-\mu \frac{1}{2} w^T w - \xi \frac{1}{2} \sum_{k=1}^N e_k^2\right)$.

After integration over all w and b values, the upper expression yields the following posterior probability for the single-class problem: $p(y | x, D) \propto p(y) p(x | y, D)$, where $p(y)$ is the prior distribution and the likelihood is given by:

$$p(x | y, D) = \frac{1}{\sqrt{2\pi(\xi^{-1} + \sigma_e^2)}} \exp\left(-\frac{m_e^2}{2(\xi^{-1} + \sigma_e^2)}\right). \tag{5.15}$$

The mean value is evaluated as:

$$\begin{aligned}
 m_e &= w^T \varphi(x) - \hat{m} = \sum_i a_i K(x, x_i) - m_d, \\
 \text{and, } \hat{m}_d &= \frac{1}{N} \sum_k a_i \sum_i K(x_k, x_i).
 \end{aligned} \tag{5.16}$$

Additionally, the variance is calculated as:

$$\sigma_e^2 = [\varphi(x) - \hat{m}]^T Q [\varphi(x) - m]. \tag{5.17}$$

The output of the anomaly detector in Eq. (5.15) is interpreted by using the probabilistic output values. If this probability value is below a certain threshold value, the data sample will be considered as belonging to the normal operating condition. Otherwise, if it exceeds that threshold

value, this instance will be thought of as belonging to the abnormal operating condition. Before actual fault detection, the anomaly detector may experience a region of uncertainty where the probabilistic output oscillates between 0 and 1. To characterize this behavior, this study activates an alarm system if the expectation exceeds a certain confidence level (i.e., a 95% probability of failure).

5.2.3 Recurrent Least-Square Support Vector Regression (Recurrent-LSSVR) for Predicting Future DD-HI Values and Estimating RUL

Once an abnormal change in the DD-HI behavior is detected, which is an indication of potential failure, the prognosis module is activated. The prognosis module attempts to model the dynamics of this impending failure with the goal of RUL estimation. As this dynamic behavior is highly non-linear, traditional models (e.g., linear, polynomial, and other higher-order regressions) are unable to appropriately model this complex behavior. Thus, in this chapter, a robust Recurrent-LSSVR framework is utilized to achieve an accurate RUL estimation result.

To define LSSVR for a given dataset $\{(x_k, y_k)\}_{k=1}^N$, with $x_k \in R^N$ and $y_k \in R$, the LSSVR optimization problem can be expressed as follows:

$$\begin{aligned} \min_{w, b, e} J(w, e) &= \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{k=1}^N e_k^2 \\ \text{s.t. } y_k &= w^T \theta(x_k) + b + e_k, \quad k = 1, \dots, N. \end{aligned} \quad (5.18)$$

Here, w is the trainable weight vector and b is the scalar bias. The Lagrangian dual formulation of Eq. (5.18) can be defined as shown below:

$$L(w, b, e; \alpha) = J(w, e) - \sum_{k=1}^N \alpha_k \{w^T \varphi(x_k) + b + e_k - y_k\}. \quad (5.19)$$

Partial derivation of Eq. (5.19) and simplification lead to the following linear solution of the proposed LSSVR:

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \Omega + \mathbf{I} / V_\gamma \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}, \quad (5.20)$$

where $\boldsymbol{\alpha}$ are Lagrangian multipliers of form $\boldsymbol{\alpha} = [\alpha_1; \dots; \alpha_N]$. Additionally, $\mathbf{y} = [y_1; \dots; y_N]$, $\mathbf{1}_v = [1; \dots; 1]$, $\Omega = \Phi(x_i)^T \Phi(x_i) = K(x_i, x_j)$, and the diagonal matrix V is given by $V_\gamma = \text{diag}\left(\left[\frac{1}{\gamma v_1}; \dots; \frac{1}{\gamma v_N}\right]\right)$. A common weighting factor v_i is determined based upon the error variable $e_i = \alpha_k / \gamma$ in subsequent training; this is described below [122]:

$$v_i = \begin{cases} 1 & \text{if } |e_i / \hat{s}| \leq d_1 \\ \frac{d_2 - |e_i / \hat{s}|}{d_2 - d_1} & \text{if } d_2 \leq |e_i / \hat{s}| \leq d_2 \\ 10^{-4} & \text{Otherwise} \end{cases} \quad (5.21)$$

Where, \hat{s} is the standard deviation of the error variable. The regression function is then given by the following formulation:

$$y(x) = \sum_{k=1}^N \alpha_k K(x, x_k) + b, \quad (5.22)$$

where α_k and b are solutions to the linear system presented in Eq. (20).

To perform a time-series data prediction, this chapter extends this formulation with a recurrent training model, as shown below and depicted in Figure 4:

$$A_N \boldsymbol{\alpha}_N = Y_N \quad (5.23)$$

$$\text{where, } Y_N = F(y_{N-1}, y_{N-2}, \dots, y_{N-q}), \quad (5.24)$$

$$\text{where, } A_{N+1} = \begin{bmatrix} A_n & a \\ a^T & c \end{bmatrix}, \boldsymbol{\alpha}_{N+1} = \begin{bmatrix} \alpha_N \\ \alpha_{N+1} \end{bmatrix}, Y_{N+1} = \begin{bmatrix} Y_N \\ Y_{N+1} \end{bmatrix}$$

$$\text{and } a = [1; \phi(x_1, x_{n+1}); \dots; \phi(x_n, x_{n+1})], \text{ and } c = C^{-1} + \phi(x_{n+1}, x_{n+1}).$$

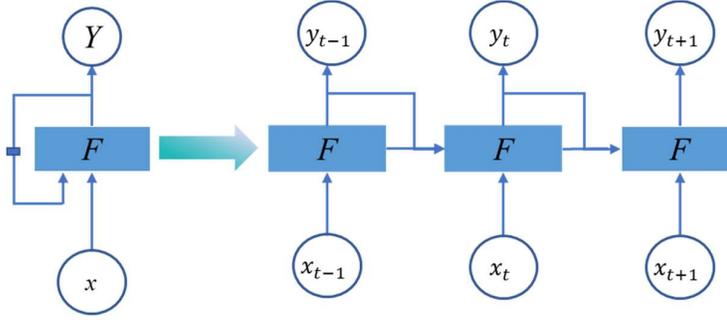


Figure 5.5: Proposed recurrent training framework for least-square support vector regression.

One can calculate the inverse A_{N+1}^{-1} , which is updated at each step, according to the Golub and Van Loan formulation [125] using the following expression:

$$A_{N+1}^{-1} = \begin{bmatrix} A_{N+1}^{-1} & 0^T \\ 0 & 0 \end{bmatrix} + [c - a^T A_N^{-1} a]^{-1} \begin{bmatrix} A_N^{-1} a \\ -1 \end{bmatrix} \times [a^T A_N^{-1} a \quad -1]. \quad (5.25)$$

Note that in the conventional formulation, the size of the matrix will grow linearly with the addition of new data instances, which makes the scheme ill-suited for online application. It also makes the algorithm inefficient for non-stationary dynamics since it requires infinite memory. To resolve these shortcomings, forward learning is used in conjunction with a decremental unlearning step, as seen in Figure 5.6. In essence, this is a pruning step that omits relatively insignificant values. Without explicit calculation of the inverse matrix A for the data point N , the following update equation is applied [126]:

$$A_{ij}^{-1} = A_{ij}^{-1} - A_{kk}^{-1} A_{ik}^{-1} A_{jk}^{-1}, \quad (5.26)$$

where $i, j = 1, \dots, N; i, j \neq k$ and A_{ij}^{-1} defines the item of the i^{th} row and j^{th} column of A_{N+1}^{-1} .

5.2.4 Failure Threshold

The failure threshold in existing studies is usually determined using a constant value, such as the average value of the HI or the highest value of the HI for all the bearings in the run-to-failure historical data [29, 127]. In this research, this chapter dynamically determines the failure threshold

based on calculating an instantaneous gradient of a linear model in the updated window. When the future values of DD-HI are estimated using Recurrent-LSSVR, an updated window is obtained in each period. A linear regression model is derived after each iteration of the estimation process for the purpose of failure threshold detection. The gradient of the obtained linear fit model is assessed against a small value on the basis of the slope in the historical data at a severe stage for a number of bearings. When this value is exceeded by a gradient value (e.g., a gradient ≥ 0.0005) of the sliding window with a size of 50 signals, the system is supposed to reach its failure threshold; this indicates that a bearing is experiencing a severe health condition. Once the failure threshold is detected, the algorithm stops estimating new DD-HI values and the RUL is obtained as the number of predicted health index values by taking into the account the sampling time between each of the instances.

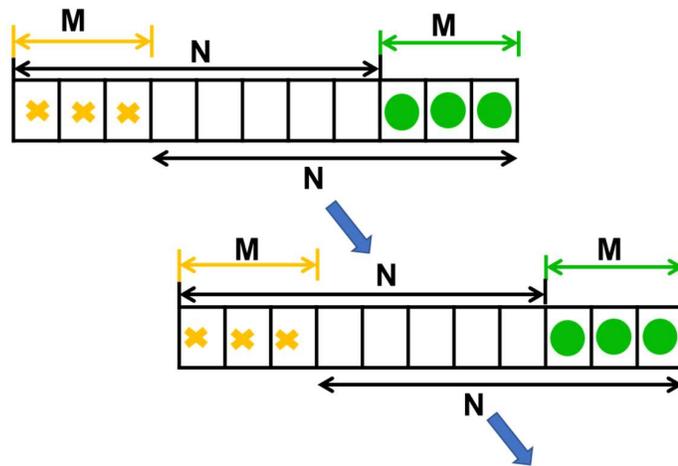


Figure 5.6: Updating process of the proposed Recurrent-LSSVR model.

5.3 Bearing test rig and Run-to-failure dataset

The proposed prognostic scheme is tested on run-to-failure REBs data of Intelligent Maintenance System (IMS) center in the University of Cincinnati [29]. These data are publicly available through the NASA Ames Prognostics Data Repository [127]. Figure 6 depicts the schematic diagram of the experimental setup that is used to collect the bearing data. Four new Rexnord double row bearings (model ZA-2115) were placed on one shaft during the experiment, as shown in Figure 6. PCB 353B33 accelerometers were mounted on each bearing housing to obtain vibration data. The data were recorded at 2000 revolution per minutes (RPM) of shaft speed and sampled at a rate of 20,000 . During the experiment, a radial load of 6000 lbs. was applied to the shaft and bearings. The data were recorded as snapshots of the vibration acceleration signal for

periods of one second. Snapshots were taken every 10 min. The IMS data used for prognosis consists of three data sets acquired during the three experiments.

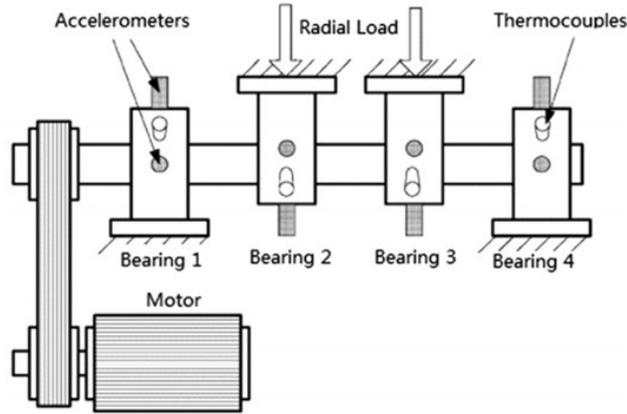


Figure 5.7: Experimental setup and component location for recording run-to-test failure data for bearing RUL estimation [29].

5.4 Results verification and discussion

The experimental results of the proposed prognosis framework for accurate anomaly detection and robust RUL estimation are presented in this section using IMS run-to-failure data of experiment no. 2 (see Section III). The data sequences corresponding to bearing no. 2 and bearing no. 4 in the second experimental data set are analyzed in this study.

As discussed in Section II, the proposed prognosis method resolves three important issues for the health management of REBs: DD-HI construction obtains an evolving trend of bearing degradation, Bayesian-OCLSSVM allows for precise detection of abnormal data instances, and the recurrent-LSSVR model predicts the future degradation trends of DD-HI for robust RUL estimation.

Figure 5.8 illustrates the comparison between the degradation trends obtained using the proposed DD-HI indicator and other widely used health indices, such as RMS and kurtosis [128], for bearings no. 2 and no. 4. From Figure 5.8, it is clearly seen that the proposed health indicator shows a reasonable evolving trend over time, while the trends obtained by the counterpart health indices are not monotonous in areas containing a high degree of oscillation.

To perform accurate detection of the abnormal health indicator behavior and define the TTS point for RUL prediction, the Bayesian-OCLSSVM anomaly detector is applied to the proposed DD-HI. The Bayesian-OCLSSVM anomaly detection module is used to discriminate between two Boolean states: the absence of fault (normal condition) and the presence of a fault

(faulty condition), based on the probability of states. The results of failure detection using Bayesian-OCLSSVM for bearings no. 2 and no. 4 are presented in Figure 5.9. As seen in Figure 5.9(a), some significant oscillations in the probabilistic output of the detector around the 600th time samples are observed. However, these oscillations do not become clear enough to ensure the existence of a fault, i.e., $(P\{DD - HI\ values\} > 0.95)$, until after the 701st time sample. The results presented in Figure 8(b) correspond to bearing no. 4 from the second IMS run-to-failure test. From Figure 5.9(b), it is clearly seen that failure was detected by Bayesian-OCLSSVM with 100% accuracy around the 601st time sample. Thus, the data samples at which the faults of bearings were detected with 100% accuracy can be considered as TTS points and further utilized for RUL estimation purposes.

As Bayesian-OCLSSVM can be used for accurate detection, the RUL is estimated by predicting future values of the DD-HI based on Recurrent-LSSVR. The initial window of faulty data used for pre-training the Recurrent-LSSVR model is further advanced to stay updated regarding the bearing health condition by including the newly predicted values of the proposed health indicator and excluding the outdated (oldest) ones. The estimation process of future values continues until the predicted value crosses the failure threshold. To determine the failure threshold, this research evaluates the linear regression model and its gradient on each updated window of data. This process continues and its gradient is updated at each step. The proposed gradient-based dynamic thresholding technique is given in Figure 5.10; here, the red color indicates the steepness of the predicted trajectory of the future DD-HI values. A steeper trajectory indicates a higher gradient, which means the RUL estimation process is approaching the failure threshold. Once this gradient value exceeds a preset value (i.e., 0.0005, as determined from historical data), then the bearing is declared as failed.

As a proof-of-concept, Figure 5.11 displays the prediction results of future DD-HI values obtained by the Recurrent-LSSVR model. The predicted curve (green curve) confirms that the prognosis scheme trained on a limited amount of data (red curve) is capable of capturing non-linearity in the degradation process. Once the predicted curve exceeds the failure threshold, the RUL is then calculated by considering the number of predicted values and the time duration between samples. The RUL can be obtained at a given time using Eq. (5.27):

$$\hat{r}(t_i) = k \cdot \Delta T. \quad (5.27)$$

Here, k corresponds to the number of estimated samples of the health indicator, while ΔT represents the sampling period of the data samples. The dataset used in this study has a sampling period of 10 min [127]. Therefore, Table I presents the estimated RUL values as an average of all estimated values in the variable size of the window (i.e., size = 10-50 instances). According to the results in Table I, the predicted RUL using Recurrent-LSSVR is close to the true RUL.

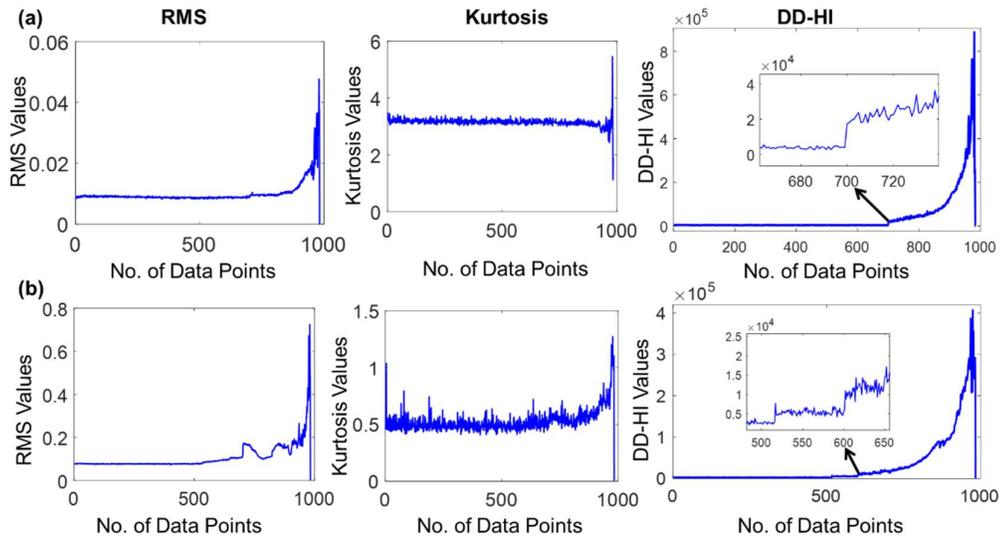


Figure 5.8: Different health indexes of bearing run-to-test failure data for (a) bearing no. 2 and (b) bearing no. 4 using RMS, Kurtosis, and DD-HI (proposed).

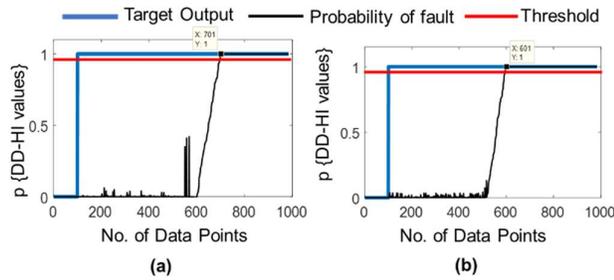


Figure 5.9: Anomaly detector output for (a) bearing no. 2 and (b) bearing no. 4. The first 50 instances of the initial data are used for pre-training the anomaly detector.

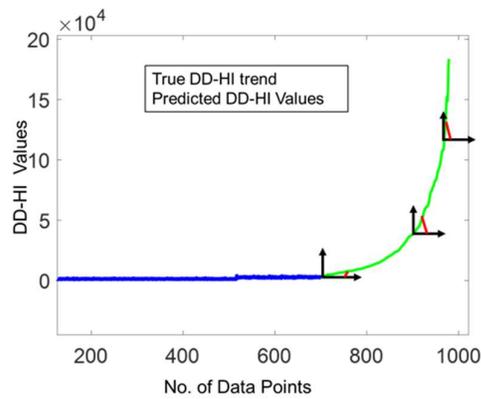


Figure 5.10: Gradient-based dynamic thresholding method at various points. In this figure, blue indicates the true trend, green indicates the predicted trend, and red indicates the gradient at different points.

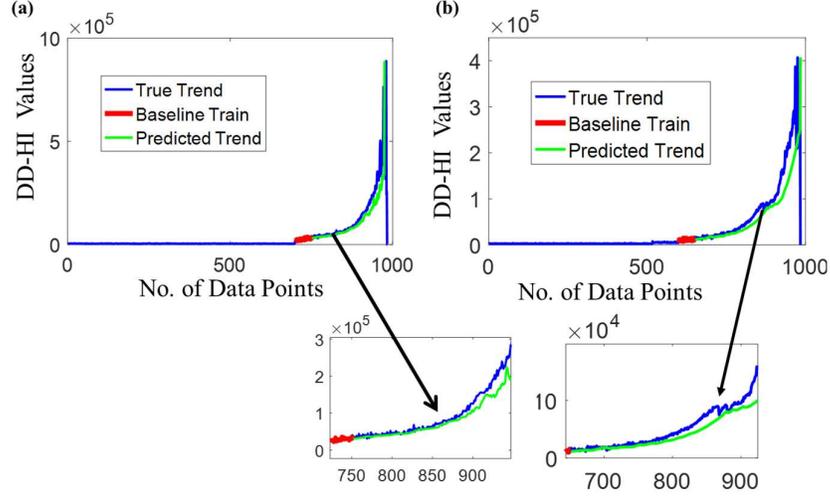


Figure 5.11: One-step-ahead prediction results of Recurrent-LSSVR for (a) bearing no. 2 and (b) bearing no. 4. After the TTS point, 50 instances are used for training to build the baseline model (red color).

Furthermore, this chapter performs various numerical analyses to verify the robustness of the proposed RUL estimation algorithm. The most widely used performance metrics are the mean-square error (MSE) and standard deviation (SD), which are defined as follows [129]:

$$MSE = \frac{1}{N} \sum_{i=1}^N [e(i)]^2, \quad (5.28)$$

where, $e(i) = r(i) - \hat{r}(i)$, N is the number of instances.

$$SD = \sqrt{\frac{\sum_{i=1}^N [e(i) - M]^2}{N-1}}, \quad (5.29)$$

here, M is the mean value of the RUL.

Table 5.1: True RUL and estimated RUL at different measurement points

Measurements	True RUL (min)	Estimated RUL (min)
1	1120	1140
2	1100	1090
3	1070	1050
4	920	950
5	800	750
6	550	620
7	270	260
8	105	100
9	70	75
10	30	20

The performance results of the proposed algorithm are compared with a couple of state-of-the-art prognostic methods, including the Paris model [114], an improved exponential model-based prognostic method [115], and an improved particle filtering-based linear regression model [130]. Table II provides the comparison results in terms of MSE and SD scores of the proposed scheme and the state-of-the-art techniques for RUL estimation. It can be seen in Table II that the proposed scheme outperforms the other techniques, yielding better MSE and SD scores.

Figure 5.12 presents the $\alpha - \lambda$ metric to validate the convergence performance for the proposed scheme. The predicted and actual RUL values are illustrated in the figure from $\lambda = 0$ to $\lambda = 1$, where 0 is the initial prediction time and 1 represents the

Table 5.3: MSE and SD scores for the proposed method and state-of-the-art Prognostics algorithms for RUL estimation

Test Set	MSE (SD) in %			
	Paris model [114]	Linear model [130]	Improved exponential model [115]	Proposed method
Bearing 2	67.3 (12.8)	84.2 (11.3)	85.1 (7.3)	94.7 (4.2)
Bearing 4	74.3 (8.6)	89.6 (9.1)	88.2 (6.5)	93.2 (3.9)

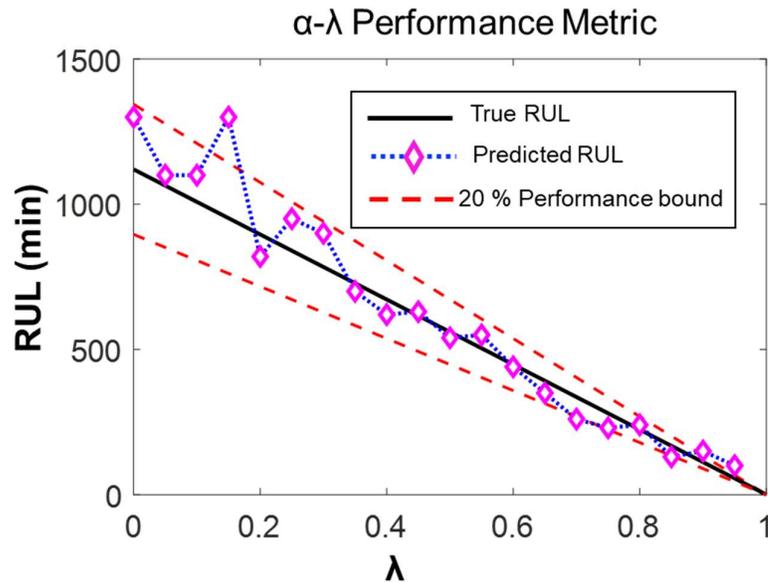


Figure 5.12: RUL prediction performance of the proposed scheme.

Table 5.4: Comparisons between proposed and EKF-based prognostic approaches [13] for RUL estimation using various HI

set	Test Measurement	Proposed Recurrent-LSSVR		EKF	
		DD-HI	RMS	RMS	Variance
Bearing 2	1	97%	91%	96%	43%
	2	92%	90%	90%	81%
	3	89%	82%	56%	68%
	4	87%	84%	64%	53%
	5	97.5%	76%	75%	49%
Average		94.7%	84.6%	76.2%	58.8%
Bearing 4	1	97%	91%	89%	54%
	2	95%	89%	91%	70%
	3	92%	82%	61%	24%
	4	86%	78%	53%	36%
	5	96%	76%	49%	42%
Average		93.2%	83.2%	68.6%	45.2%

end-of-life (EOL) of the bearings. The tolerance boundary for the RUL estimates is 20% of the actual RUL values. Figure 5.12 clearly shows that almost all of the estimated values are inside the acceptable tolerance region, i.e., $[(1 - \alpha)r(t_i)] \leq \hat{r}(t_i) \leq [(1 + \alpha)r(t_i)]$, where $r(t_i)$ corresponds to the actual RUL and $\hat{r}(t_i)$ represents the predicted RUL.

Singleton et al. [13] utilized a prognostic framework that applied an extended Kalman filter (EKF) while using the time-domain statistical parameters namely RMS and variance as health indexes. One of the motivations of this study was to develop an improved health index, e.g., DD-HI; therefore, the proposed method using DD-HI is compared with the method using the RMS and variance. In [13], the performance was measured based on the percentage values of the estimated RUL that reside within 20% of the actual RUL. The performance of the proposed schemes is compared with the EKF approach using the performance metrics provided in [9]. The proposed Recurrent-LSSVR-based scheme is also applied to estimate the RUL of REBs using the same health index like the one used in [13], i.e., the RMS values of run-to-failure data. The experimental results are given in Table 5.4. The experimental results shown in Table III, which represent an average of five different measurements, demonstrate that the proposed methodology outperforms the EKF-

based method across all the test sets. Moreover, the proposed health index, i.e., DD-HI, demonstrated improved prognostic performance compared to the other health indicators, i.e., the RMS and variance.

5.5 Conclusion

This study strengthened the PHM capabilities for robust RUL estimation of REBs by developing a solution containing an improved health index and two important modules: a novel anomaly detector for TTS detection and a robust regression model for predicting the RUL. Specifically, the authors proposed a frequency-domain health index, referred to as DD-HI, by defining narrow-band degree-of-defectiveness regions around bearing defect frequencies. Then, Bayesian-OCLSSVM was utilized as a probabilistic anomaly detector to detect TTS for RUL estimation. Additionally, a robust regression model, referred to as Recurrent-LSSVR, was applied in an online manner to predict future values of the REB degradation trend shown by DD-HI. The proposed data-driven prognosis framework was verified using a publicly available run-to-failure dataset. The experimental results validate the efficacy and robustness of the proposed framework in terms of MSE and SD metrics and the convergence compared with recent cutting-edge prognostics methods.

Chapter 6

Summary of Contributions and Future Work

6.1 Summary of Contributions

This dissertation presents a reliable and robust data-driven framework for rotating machines fault diagnosis and failure prognosis using advanced signal processing and machine learning techniques. This framework works in tandem to provide a complete PHM solution capable of operating in nonlinear and non-stationary environments and includes robust condition-based monitoring (CBM), reliable diagnostics, and true failure prognostics schemes.

In Chapter 2, this dissertation develops a CBM algorithm in operation based on time-frequency based envelope analysis and optimum sub-band selection. In practice CMB data (i.e., signal) about machine health are massive, it is important to select an information portion of the signal to analyze for defecting impending failures. Therefore, this chapter performs WPT-EA to generate a series of sub-band signal and then each sub-band signal is quantified using the proposed evaluation metrics— gaussian mixture model-based degree-of-defectives ratio (DDR). This evaluation metrics, DDR, is highly effective since it takes the ratio between the defect components and residual components and considers an actual dynamic of the failure mechanism. To verify the effectiveness, a 2D tool is developed as a percent of DDR values. The results indicate that proposed (WPT-EA + DDR) is highly cable of selecting a narrow-band signal from a stream of signal and outperforms state-of-the-art.

Chapter 3 and Chapter 4 are devoted to reliable fault diagnosis for identifying the root cause of failure, such as types, location, sizes, in various operating scenarios. In Chapter 3, this dissertation proposes a reliable fault diagnosis scheme using heterogenous feature extraction models (HFEMs) and hybrid feature selection (HFS) scheme. HFEMs are comprised of various signal process techniques namely time-domain, frequency-domain and time-frequency domain analysis to extract as many as possible features to represent fault condition distinctly. Statistical

parameters such as RMS, kurtosis, skewness, impulses, etc. of time-domain and corresponding frequency-domain are well verified to represent faults. In addition, wavelet packet transforms (WPT) is applied that simultaneously analyzes a measured signal concerning the fault in both the time and frequency domains where impulse information is detectable. A three-level WPT is used for this study and in the final level, 8 nodes (e.g., sub-bands) are generated. For each node, relative wavelet energy (RWE) information is calculated as useful information regarding the fault. Thus, HFEMs ensure a robust feature extraction process but high-dimensional since it considers as many as possible features regarding faults. In practices, a feature extraction, such as HFEMs, involving high-dimensional feature vectors are either redundant or irrelevant to a classifier (e.g., SVM), which may be a reason for degrading diagnostics performance. Therefore, this chapter then applies the HFS technique for selecting the most discriminant feature subset for improving diagnostics performance. This HFA combines both filter and wrapper approaches and applies sequential floating forward search (SFFS) with novel objective function for selecting the most discriminative features subset. The efficacy of the proposed HFS methodology is verified with a fault diagnosis application for low-speed bearing for identifying the various single and multiple-combined fault at various scenarios. The results indicated that the proposed method is more effective for identifying the most discriminatory feature subset by achieving diagnostics performance improvement of 4.16% to 19.0% in the average classification accuracy.

Fault diagnosis in variable operating conditions (such as rotational speed) is still a challenging problem since the fault characteristics are significantly changed because of varied operating conditions. To deal with this issue, in Chapter 4, this dissertation developed a highly efficient bearing fault diagnosis scheme under variable speed conditions using an acoustic spectrum imaging (ASI) of acoustic emission (AE) signals as a precise health state and an adaptive deep convolutional neural network (ADCNN). In ASI, the amplitudes of the spectral components of the windowed time-domain acoustic emission signal are transformed into spectrum imaging. ASI provides a visual representation of acoustic emission spectral features in images. This ensures enhanced spectral images for ADCNN testing and training, and thus provides a robust classifier technique with high diagnostic accuracy. This study validated our proposed method (ASI + ADCNN) by using the health images of four different health conditions for three different rotational speeds. The proposed fault diagnosis architecture yields an average accuracy of 92.67%, which establishes this proposed method as invariant to variations of shaft speed. In addition, the proposed framework outperforms three the state-of-the-arts, yielding an overall performance improvement of 16.55%, 1.15%, and 20.21% in all scenarios respectively.

Thirdly, a novel approach is proposed for the health prognosis of rolling element bearings in Chapter 4, which infers the health of a bearing through a dimensionless health indicator (HI) and estimates its RUL using dynamic regression models. The proposed HI measures the instantaneous vibration level of the bearing with respect to a normal baseline value. The regression models are recursively updated to capture the evolving trend in the bearings HI and are then used to project the future values of the health indicator and estimate the RUL of the bearing. The RUL of a bearing is estimated after determining the time to start prediction (TSP) using a new alarm bound technique

Chapter 5 strengthened the PHM capabilities for robust RUL estimation of REBs by developing a solution containing an improved health index and two important modules: a novel anomaly detector for TTS detection and a robust regression model for predicting the RUL. Specifically, the authors proposed a frequency-domain health index, referred to as DD-HI, by defining narrow-band degree-of-defectiveness regions around bearing defect frequencies. Then, Bayesian-OCLSSVM was utilized as a probabilistic anomaly detector to detect TTS for RUL estimation. Additionally, a robust regression model, referred to as Recurrent-LSSVR, was applied in an online manner to predict future values of the REB degradation trend shown by DD-HI. The proposed data-driven prognosis framework was verified using a publicly available run-to-failure dataset. The experimental results validate the efficacy and robustness of the proposed framework in terms of MSE and SD metrics and the convergence compared with recent cutting-edge prognostics methods.

6.2 Future Work

This dissertation main efforts to fortified PHM capabilities to ensure a robust health monitoring system for massive industrial cases. Moreover, avenues for further research in the future are suggested.

Online condition monitoring (OCM) system is still a challenging issue since it requires real-time data processing and a continuous update process for efficient memory management. For real-world applications, as rotating machines operate for several years, therefore, it is important to design an efficient OCM system for up-to-date information about machine operating conditions.

For reliable fault diagnosis, feature engineering and feature learning schemes are mostly used. In feature engineering, this dissertation first extracts as many as possible features to represent fault condition, and then apply a feature selection algorithm to select the best feature subset. There are various feature selection algorithms such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Simulated Annealing (SA), Quadratic Programming Feature Selection (QPFS), Information Theory Based Feature Selection (ITFS) can be investigated to attain near

global performance. In addition to the state-of-the-art feature selection algorithms, feature selection techniques need an innovative object (or fitness) function. This objective function is problem depended and mostly found effective in Gaussian distribution cases. So, it is also important to consider non-Gaussian distributions since practical applications are mostly composed of nonlinear and non-Gaussian system. For novel objective function design, structural analysis and nonlinear properties of distributions can be adopted.

Recently, feature learning (i.e., deep learning) processes are gaining popularity in fault diagnosis application since it automates feature engineering and classification processes from raw data. In this process, if the data are not validated, the diagnostics performance can be guaranteed. It is important to ensure appropriate sensors and sensing strategies to acquire relevant data that are promising to improve substantially the reliability, cost-effectiveness, coverage, and sensitivity of the fault signatures of the monitoring and fault tracking devices. Furthermore, data characteristics of faults are also significantly varied with operating conditions namely fault types, operating speed, load, fault severities, and noises. Therefore, the consideration of various operating factors could be a viable solution for the future condition-based monitoring system.

Prognosis is the most promising feature and less explored in the PHM research filed in compare with fault diagnosis and fault detection-based condition monitoring system. Prognosis is the process of RUL estimation. RUL estimation always incurred with several challenges for examples, machine health degradation model and failure threshold to define end-of-life (EOF). Degradation models can be model-based and data-driven. The fusion of model-based and data-driven could be the ultimate vision of PHM capabilities that could able to utilize advantages of both model-based and data-driven systems for the mission-critical solution.

Publications

International Journals

1. Junayed Hasan, **M M Manjurul Islam**, Jong-Myon Kim, “Acoustic spectral imaging and transfer learning for reliable bearing fault diagnosis under variable speed conditions,” *Measurement*, Vol. 106, February 2019, pp620-631, DOI: 10.1016/j.measurement.2019.02.075
2. **M M Manjurul Islam**, Jong-Myon Kim, “Automated bearing fault diagnosis scheme using 2D representation of wavelet packet transform and deep convolutional neural network,” *Computer in Industry*, Vol. 106, April 2019, pp142-153, DOI: 10.1016/j.compind.2019.01.008
3. Muhammad Sohaib, **M M Manjurul Islam**, Jaeyoung Kim, Jong-Myon Kim, “Leakage detection of a spherical water storage tank in a chemical industry using acoustic Emissions,” *Applied Sciences*, Vol. 9, No. 1, January 2019, pp196, DOI:10.3390/app9010196
4. **M M Manjurul Islam**, Muhammad Sohaib, Jaeyoung Kim, Jong-Myon Kim, “Crack classification of a pressure vessel using feature selection and deep learning methods,” *Sensors (MDPI)*, Vol. 18, No. 12, December 2018, pp4379, DOI:10.3390/s18124379.
5. Alexander Prosvirin, **M M Manjurul Islam**, Jaeyoung Kim, Jong-Myon Kim, “Rub-impact fault diagnosis using an effective IMF selection technique in ensemble empirical mode decomposition and hybrid feature models,” *Sensors (MDPI)*, Vol. 18, No. 7, June 2018, pp2040, DOI:10.3390/s18072040.
6. **M M Manjurul Islam**, Jong-Myon Kim, “Reliable multiple combined fault diagnosis of bearings using heterogeneous feature models and multiclass support vector machines,” *Reliability Engineering & System Safety*, Vol.184, February 2018, pp55-66, DOI:10.1016/j.ress.2018.02.012.
7. Wasim Ahmad, Sheraz Khan, **M M Manjurul Islam**, Jong-Myon Kim, “A reliable technique for remaining useful life estimation of rolling element bearings using dynamic regression models,” *Reliability Engineering & System Safety*, Vol.184, February 2018, pp67-76, DOI: 10.1016/j.ress.2018.02.003.

8. **M M Manjurul Islam**, Jong-Myon Kim, “Time–frequency envelope analysis-based subband selection and probabilistic support vector machines for multi-fault diagnosis of lowspeed bearings,” *Journal of Ambient Intelligence and Humanized*, June 2017, pp1-16, DOI:10.1007/s12652-017-0585-2.
9. **M M Manjurul Islam**, Jong-Myon Kim, “Reliable bearing fault diagnosis using Bayesian inference-based multi-class support vector machines,” *The Journal of the Acoustical Society of America Express Letter*, April 2016, Vol. 141, No. 2, February 2017, pp89-95, DOI:10.1121/1.4976038.

Under review

1. Alexander Prosvirin, **M M Manjurul Islam**, Jong-Myon Kim, “An Improved Algorithm for Selecting IMF Components Using Ensemble Empirical Mode Decomposition for Rub-Impact Fault Diagnosis,” (**Under review**).
2. **M M Manjurul Islam**, Alexander Prosvirin, Jong-Myon Kim, “Data-driven prognostic scheme for rolling-element bearings using a new health index and variants of least-square support vector machines,” (**Under review**).

Book Chapters

1. Se-Jong Kang, Jae-Young Kim, In-Kyu Jeong, Kichang Im, **M M Manjurul Islam**, Jong-Myon Kim, “An Improved Gas Classification Technique Using New Features and Support Vector Machines,” Presented In: Tenth International Conference on Soft Computing and Pattern Recognition, SoCPaR 2018, December 2018, Porto, Portugal, Chapter: *Lecture Notes in Intelligent Systems and Computing, Springer*, Vol. 942, pp158- 166, DOI:10.1007/978-3-030-17065-3_16.
2. **M M Manjurul Islam**, Alexander Prosvirin, Jong-Myon Kim, “Intelligent Rub-impact fault diagnosis based on genetic algorithm-based IMF selection in ensemble empirical mode decomposition and diverse features models,” Presented In: 19th Intelligent Data Engineering and Automated Learning, IDEAL 2018, November 2018, Madrid, Spain, Chapter: *Lecture Notes in Computer Science, Springer*, Vol. 11314, pp147-155, DOI:10.1007/978-3-319-89656-4_12.
3. **M M Manjurul Islam**, Jong-Myon Kim, “Motor bearing fault diagnosis using deep convolutional neural networks with 2D analysis of vibration signal,” Presented In: 31st

- Canadian Conference on Artificial Intelligence, Canadian AI 2018, May 2018, Toronto, Canada, Chapter: *Lecture Notes in Artificial Intelligence, Springer*, Vol. 10832, pp144-155, DOI:10.1007/978-3-030-03493-1_16.
4. Farzin Piltan, **M M Manjurul Islam**, Jong-Myon Kim, "Input-output fault diagnosis in robot manipulator using fuzzy LMI-tuned PI feedback linearization observer based on nonlinear intelligent ARX model," Presented In: 2nd International Conference on Computer, Communication and Computational Sciences, IC4S 2017, October 2017, Phuket, Thailand, Chapter: *Lecture Notes in Computer Science, Springer*, Vol. 759, pp305-315, DOI:10.1007/978-981-13-0341-8_28.
 5. **M M Manjurul Islam**, Rashedul Islam Jong-Myon Kim, "A hybrid feature selection scheme based on local compactness and global separability for improving roller bearing diagnostic performance," Presented In: 3rd Australasian Conference on Artificial Life and Computational Intelligence, ACALCI 2017, January 2017, Melbourne, Australia, Chapter: *Lecture Notes in Computer Science, Springer*, Vol. 10142, pp180-192, DOI:10.1007/978-3-319-51691-2_16.
 6. **M M Manjurul Islam**, Sheraz Khan, Jong-Myon Kim, "Multi-fault diagnosis of roller bearings using support vector machines with an improved decision strategy," Presented In: 11th International Conference on Intelligent Computing, ICIC 2015, August 2015, Fuzhou, China, Chapter: *Lecture Notes in Artificial Intelligence, Springer*, Vol. 9227, pp538-550, DOI:10.1007/978-3-319-51691-2_16. (Best Paper Award)

International Conferences

1. **M M Manjurul Islam**, Alexander Prosvirin, Jong-Myon Kim, Robust condition monitoring of rub-impact faults in rotor system using empirical mode decomposition with IMF selection, The 2017 Engineering and Arts Society Conference in Korea (EASKO 2017), December 15-16, 2017, Ulsan, South Korea. **(Best Paper Award)**
2. Alexander Prosvirin, **M M Manjurul Islam**, Jong-Myon Kim, Fault prediction of rolling element bearings using one class least squares SVM, The 2017 Engineering and Arts Society Conference in Korea (EASKO 2017), December 15-16, 2017, Ulsan, South Korea. **(Best Paper Award)**
3. Young-Hun Kim, **M M Manjurul Islam**, Md Rashedul Islam, Jong-Myon Kim, Genetic algorithm based discriminant feature selection for improved fault diagnosis of induction motor, 19-th International Conference on Artificial Intelligence, (ICAI 2017), July 17-20, 2017, Las Vegas, Nevada, USA.

-
4. Jeong Inkyu, Dileep Kumar, **M M Manjurul Islam**, Cheol-Hong Kim, and Jong-Myon Kim, Optimized mobile robotic navigation based on fuzzy logic control, 19-th International Conference on Artificial Intelligence, (ICAI 2017), July 17-20, 2017, Las Vegas, Nevada, USA.
 5. **M M Manjurul Islam**, Jong-Myon Kim, Reliable faults diagnosis of roller bearings using hybrid feature models and improved multiclass support vector machines with classifiers discriminant analysis, Asia Pacific Conference of the Prognostics and Health Management Society (PHMAP 2017), July 12-15, 2017, Jeju, Korea. **(Best Paper Award)**
 6. **M M Manjurul Islam**, Jong-Myon Kim, "Envelope analysis-based sub-band selection and improved support vector machines for reliable bearing multi-fault diagnosis, The 11th International Conference on Multimedia and Ubiquitous Engineering (MUE 2017), May 22-24, 2017, Seoul, South Korea. **(Best Paper Award)**
 7. **M M Manjurul Islam**, Jong-myon Kim, Fault detection and diagnosis for rolling element bearing using wavelet analysis and support vector machine, The 2016 Engineering and Arts Society Conference in Korea (EASKO 2016), November 10-11, 2016, Ulsan, South Korea.

References

- [1] M. Kang and M. G. Pecht, "Machine Learning: Diagnostics and Prognostics," in *Prognostics and Health Management of Electronics*, 2019.
- [2] G. J. Vachtsevanos, F. Lewis, A. Hess, and B. Wu, *Intelligent fault diagnosis and prognosis for engineering systems*. Wiley Hoboken, 2006.
- [3] M. Kang, J. Kim, L. M. Wills, and J. M. Kim, "Time-Varying and Multiresolution Envelope Analysis and Discriminative Feature Analysis for Bearing Fault Diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 12, pp. 7749-7761, 2015.
- [4] M. Kang, M. R. Islam, J. Kim, J. M. Kim, and M. Pecht, "A Hybrid Feature Selection Scheme for Reducing Diagnostic Performance Deterioration Caused by Outliers in Data-Driven Diagnostics," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3299-3310, 2016.
- [5] X. Yuan, Y. Wang, C. Yang, Z. Ge, Z. Song, and W. Gui, "Weighted Linear Dynamic System for Feature Representation and Soft Sensor Application in Nonlinear Dynamic Industrial Processes," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1508-1517, 2018.
- [6] Z. Lou, D. Shen, and Y. Wang, "Two-step principal component analysis for dynamic processes monitoring," *The Canadian Journal of Chemical Engineering*, vol. 96, no. 1, pp. 160-170, 2018.
- [7] K. E. S. Pilario and Y. Cao, "Canonical Variate Dissimilarity Analysis for Process Incipient Fault Detection," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 12, pp. 5308-5315, 2018.
- [8] Z. Li, H. Fang, M. Huang, Y. Wei, and L. Zhang, "Data-driven bearing fault identification using improved hidden Markov model and self-organizing map," *Computers & Industrial Engineering*, vol. 116, pp. 37-46, 2018/02/01/ 2018.
- [9] M. Islam, M. Sohaib, J. Kim, and J.-M. Kim, "Crack Classification of a Pressure Vessel Using Feature Selection and Deep Learning Methods," *Sensors*, vol. 18, no. 12, p. 4379, 2018.

-
- [10] M. M. M. Islam and J.-M. Kim, "Time–frequency envelope analysis-based sub-band selection and probabilistic support vector machines for multi-fault diagnosis of low-speed bearings," *Journal of Ambient Intelligence and Humanized Computing*, 2017/10/09 2017.
- [11] M. M. Manjurul Islam and J.-M. Kim, "Reliable multiple combined fault diagnosis of bearings using heterogeneous feature models and multiclass support vector Machines," *Reliability Engineering & System Safety*, 2018/02/06/ 2018.
- [12] L. Saidi, J. Ben Ali, M. Benbouzid, and E. Bechhofer, "An integrated wind turbine failures prognostic approach implementing Kalman smoother with confidence bounds," *Applied Acoustics*, vol. 138, pp. 199-208, 2018/09/01/ 2018.
- [13] R. K. Singleton, E. G. Strangas, and S. Aviyente, "Extended Kalman Filtering for Remaining-Useful-Life Estimation of Bearings," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 3, pp. 1781-1790, 2015.
- [14] I.-K. Jeong, M. Kang, J. Kim, J.-M. Kim, J.-M. Ha, and B.-K. Choi, "Enhanced DET-Based Fault Signature Analysis for Reliable Diagnosis of Single and Multiple-Combined Bearing Defects," *Shock and Vibration*, vol. 2015, p. 10, 2015, Art. no. 814650.
- [15] M. Kang, J. Kim, B.-K. Choi, and J.-M. Kim, "Envelope analysis with a genetic algorithm-based adaptive filter bank for bearing fault detection," *The Journal of the Acoustical Society of America*, vol. 138, no. 1, pp. EL65-EL70, 2015.
- [16] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 485-520, 2011/02/01/ 2011.
- [17] S. T. Kandukuri, A. Klausen, H. R. Karimi, and K. G. Robbersmyr, "A review of diagnostics and prognostics of low-speed machinery towards wind turbine farm-level health management," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 697-708, 1// 2016.
- [18] D. Wang, K. Tsui, and Q. Miao, "Prognostics and Health Management: A Review of Vibration Based Bearing and Gear Health Indicators," *IEEE Access*, vol. 6, pp. 665-676, 2018.
- [19] J. Wang, F. Cheng, W. Qiao, and L. Qu, "Multiscale Filtering Reconstruction for Wind Turbine Gearbox Fault Diagnosis Under Varying-Speed and Noisy Conditions," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 5, pp. 4268-4278, 2018.
- [20] A. Prosvirin, M. Islam, J. Kim, and J.-M. Kim, "Rub-Impact Fault Diagnosis Using an Effective IMF Selection Technique in Ensemble Empirical Mode Decomposition and Hybrid Feature Models," *Sensors*, vol. 18, no. 7, p. 2040, 2018.

-
- [21] D. F. Garcia *et al.*, "Spacecraft Health Monitoring Using a Biomimetic Fault Diagnosis Scheme," *Journal of Aerospace Information Systems*, vol. 15, no. 7, pp. 396-413, 2018.
- [22] Q. Shen, C. Yue, C. H. Goh, and D. Wang, "Active Fault-Tolerant Control System Design for Spacecraft Attitude Maneuvers with Actuator Saturation and Faults," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 5, pp. 3763-3772, 2019.
- [23] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483-1510, 2006/10/01/ 2006.
- [24] 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.
- [25] M. Kang, J. Kim, J. M. Kim, A. C. C. Tan, E. Y. Kim, and B. K. Choi, "Reliable Fault Diagnosis for Low-Speed Bearings Using Individually Trained Support Vector Machines With Kernel Discriminative Feature Analysis," *IEEE Transactions on Power Electronics*, vol. 30, no. 5, pp. 2786-2797, 2015.
- [26] S. A. Abdusslam, "Detection and Diagnosis of Rolling Element Bearing Faults Using Time Encoded Signal Processing and Recognition," Doctoral, School of Computing and Engineering, University of Huddersfield, 2012.
- [27] R. Tiwari and R. M. Chandran, "Multitude of objectives based optimum designs of cylindrical roller bearings with evolutionary methods," *Journal of Tribology*, vol. 137, no. 4, p. 041504, 2015.
- [28] M. M. M. Islam and J.-M. Kim, "Motor Bearing Fault Diagnosis Using Deep Convolutional Neural Networks with 2D Analysis of Vibration Signal," in *Advances in Artificial Intelligence*, Cham, 2018, pp. 144-155: Springer International Publishing.
- [29] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of Sound and Vibration*, vol. 289, no. 4, pp. 1066-1090, 2006/02/07/ 2006.
- [30] I. V. de Bessa, R. M. Palhares, M. F. S. V. D'Angelo, and J. E. Chaves Filho, "Data-driven fault detection and isolation scheme for a wind turbine benchmark," *Renewable Energy*, vol. 87, Part 1, pp. 634-645, 3// 2016.
- [31] M. Hamadache, D. Lee, and K. C. Veluvolu, "Rotor Speed-Based Bearing Fault Diagnosis (RSB-BFD) Under Variable Speed and Constant Load," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6486-6495, 2015.

- [32] G. Dong, J. Chen, and F. Zhao, "A frequency-shifted bispectrum for rolling element bearing diagnosis," *Journal of Sound and Vibration*, vol. 339, pp. 396-418, 3/17/ 2015.
- [33] X. Chen *et al.*, "Fault diagnosis based on dependent feature vector and probability neural network for rolling element bearings," *Applied Mathematics and Computation*, vol. 247, pp. 835-847, 11/15/ 2014.
- [34] Y. He and X. Zhang, "Approximate Entropy Analysis of the Acoustic Emission From Defects in Rolling Element Bearings," *Journal of Vibration and Acoustics*, vol. 134, no. 6, pp. 061012-061012, 2012.
- [35] R. Li and D. He, "Rotational Machine Health Monitoring and Fault Detection Using EMD-Based Acoustic Emission Feature Quantification," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 4, pp. 990-1001, 2012.
- [36] M. M. M. Islam, S. A. Khan, and J.-M. Kim, "Reliable bearing fault diagnosis using Bayesian inference-based multi-class support vector machines," *The Journal of the Acoustical Society of America*, vol. 141, no. 2, pp. EL89-EL95, 2017/02/01 2017.
- [37] S. Lalani and D. Doye, "Discrete Wavelet Transform and a Singular Value Decomposition Technique for Watermarking Based on an Adaptive Fuzzy Inference System," *Journal of information processing systems*, vol. 13, no. 2, pp. 340-347, 2017.
- [38] D. Wang, P. W. Tse, and K. L. Tsui, "An enhanced Kurtogram method for fault diagnosis of rolling element bearings," *Mechanical Systems and Signal Processing*, vol. 35, no. 1-2, pp. 176-199, 2// 2013.
- [39] D. H. Pandya, S. H. Upadhyay, and S. P. Harsha, "Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN," *Expert Systems with Applications*, vol. 40, no. 10, pp. 4137-4145, 2013/08/01/ 2013.
- [40] S. A. Niknam, V. Songmene, and Y. H. J. Au, "The use of acoustic emission information to distinguish between dry and lubricated rolling element bearings in low-speed rotating machines," *The International Journal of Advanced Manufacturing Technology*, vol. 69, no. 9, pp. 2679-2689, 2013/12/01 2013.
- [41] W. Caesarendra, P. B. Kosasih, A. K. Tieu, C. A. S. Moodie, and B.-K. Choi, "Condition monitoring of naturally damaged slow speed slewing bearing based on ensemble empirical mode decomposition," *Journal of Mechanical Science and Technology*, vol. 27, no. 8, pp. 2253-2262, 2013/08/01 2013.
- [42] M. Elforjani and D. Mba, "Condition Monitoring of Slow-Speed Shafts and Bearings with Acoustic Emission," *Strain*, vol. 47, no. s2, pp. 350-363, 2011/12/01 2011.

-
- [43] M. Kang, J. Kim, and J. M. Kim, "High-Performance and Energy-Efficient Fault Diagnosis Using Effective Envelope Analysis and Denoising on a General-Purpose Graphics Processing Unit," *IEEE Transactions on Power Electronics*, vol. 30, no. 5, pp. 2763-2776, 2015.
- [44] A. Prosvirin, J. Kim, and J.-M. Kim, "Bearing Fault Diagnosis Based on Convolutional Neural Networks with Kurtogram Representation of Acoustic Emission Signals," in *Advances in Computer Science and Ubiquitous Computing*, Singapore, 2018, pp. 21-26: Springer Singapore.
- [45] I.-K. Jeong, M. Kang, J. Kim, J.-M. Kim, J.-M. Ha, and B.-K. Choi, "Enhanced DET-based fault signature analysis for reliable diagnosis of single and multiple-combined bearing defects," *Shock and Vibration*, vol. 2015, 2015.
- [46] A. Widodo *et al.*, "Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine," *Expert Systems with Applications*, vol. 36, no. 3, pp. 7252-7261, 2009/04/01/ 2009.
- [47] J. Seshadrinath, B. Singh, and B. K. Panigrahi, "Vibration Analysis Based Interturn Fault Diagnosis in Induction Machines," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 340-350, 2014.
- [48] M. J. Gómez, C. Castejón, and J. C. García-Prada, "Automatic condition monitoring system for crack detection in rotating machinery," *Reliability Engineering & System Safety*, vol. 152, pp. 239-247, 2016/08/01/ 2016.
- [49] Z. Yin and J. Hou, "Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes," *Neurocomputing*, vol. 174, pp. 643-650, 2016/01/22/ 2016.
- [50] S. Huang, K. K. Tan, and T. H. Lee, "Fault Diagnosis and Fault-Tolerant Control in Linear Drives Using the Kalman Filter," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 11, pp. 4285-4292, 2012.
- [51] 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.
- [52] H. Shao, H. Jiang, X. Li, and T. Liang, "Rolling bearing fault detection using continuous deep belief network with locally linear embedding," *Computers in Industry*, vol. 96, pp. 27-39, 2018/04/01/ 2018.

- [53] E. C. Lau and H. Ngan, "Detection of motor bearing outer raceway defect by wavelet packet transformed motor current signature analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 10, pp. 2683-2690, 2010.
- [54] A. Widodo, B.-S. Yang, E. Y. Kim, A. C. C. Tan, and J. Mathew, "Fault diagnosis of low speed bearing based on acoustic emission signal and multi-class relevance vector machine," *Nondestructive Testing and Evaluation*, vol. 24, no. 4, pp. 313-328, 2009/12/01 2009.
- [55] A. Sadeghian, Z. Ye, and B. Wu, "Online Detection of Broken Rotor Bars in Induction Motors by Wavelet Packet Decomposition and Artificial Neural Networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 7, pp. 2253-2263, 2009.
- [56] M. Grasso, S. Chatterton, P. Pennacchi, and B. M. Colosimo, "A data-driven method to enhance vibration signal decomposition for rolling bearing fault analysis," *Mechanical Systems and Signal Processing*, vol. 81, pp. 126-147, 2016/12/15/ 2016.
- [57] Z. Li, X. Yan, Z. Tian, C. Yuan, Z. Peng, and L. Li, "Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis," *Measurement*, vol. 46, no. 1, pp. 259-271, 1// 2013.
- [58] J. Shen, H. Chang, and Y. Li, "Pressure vessel state investigation based upon the least squares support vector machine," *Mathematical and Computer Modelling*, vol. 54, no. 3, pp. 883-887, 2011/08/01/ 2011.
- [59] N. Ueno, C. Xu, and S. Watanabe, "Fatigue crack detection of CFRP composite pressure vessel using mechanoluminescent sensor," in *SENSORS, 2013 IEEE*, 2013, pp. 1-4.
- [60] C. Liu, D. Jiang, and W. Yang, "Global geometric similarity scheme for feature selection in fault diagnosis," *Expert Systems with Applications*, vol. 41, no. 8, pp. 3585-3595, 2014/06/15/ 2014.
- [61] Y. Yang, Y. Liao, G. Meng, and J. Lee, "A hybrid feature selection scheme for unsupervised learning and its application in bearing fault diagnosis," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11311-11320, 9// 2011.
- [62] K. Zhang, Y. Li, P. Scarf, and A. Ball, "Feature selection for high-dimensional machinery fault diagnosis data using multiple models and Radial Basis Function networks," *Neurocomputing*, vol. 74, no. 17, pp. 2941-2952, 10// 2011.
- [63] T. W. Rauber, F. d. A. Boldt, F. M. Varej, and x00E, "Heterogeneous Feature Models and Feature Selection Applied to Bearing Fault Diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, pp. 637-646, 2015.

- [64] L. Lu, J. Yan, and C. W. de Silva, "Dominant feature selection for the fault diagnosis of rotary machines using modified genetic algorithm and empirical mode decomposition," *Journal of Sound and Vibration*, vol. 344, pp. 464-483, 5/26/ 2015.
- [65] H. R. Kanan and K. Faez, "GA-based optimal selection of PZMI features for face recognition," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 706-715, 11/15/ 2008.
- [66] H. Chih-Wei and L. Chih-Jen, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415-425, 2002.
- [67] M. M. M. Islam, S. A. Khan, and J.-M. Kim, "Multi-fault Diagnosis of Roller Bearings Using Support Vector Machines with an Improved Decision Strategy," in *Advanced Intelligent Computing Theories and Applications: 11th International Conference, ICIC 2015, Fuzhou, China, August 20-23, 2015. Proceedings, Part III*, D.-S. Huang and K. Han, Eds. Cham: Springer International Publishing, 2015, pp. 538-550.
- [68] S. Zou, F. Yan, G. Yang, and W. Sun, "The Identification of the Deformation Stage of a Metal Specimen Based on Acoustic Emission Data Analysis," *Sensors*, vol. 17, no. 4, p. 789, 2017.
- [69] L. Bornn, C. R. Farrar, G. Park, and K. Farinholt, "Structural Health Monitoring With Autoregressive Support Vector Machines," *Journal of Vibration and Acoustics*, vol. 131, no. 2, pp. 021004-021004-9, 2009.
- [70] I. Aydin, M. Karakose, and E. Akin, "A multi-objective artificial immune algorithm for parameter optimization in support vector machine," *Appl. Soft Comput.*, vol. 11, no. 1, pp. 120-129, 2011.
- [71] J. D. Rodriguez, A. Perez, and J. A. Lozano, "Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 569-575, 2010.
- [72] S. A. Khan and J.-M. Kim, "Automated bearing fault diagnosis using 2d analysis of vibration acceleration signals under variable speed conditions," *Shock and Vibration*, vol. 2016, 2016.
- [73] R. Saidur, "A review on electrical motors energy use and energy savings," *Renewable and sustainable energy reviews*, vol. 14, no. 3, pp. 877-898, 2010.
- [74] I. El-Thalji and E. Jantunen, "Dynamic modelling of wear evolution in rolling bearings," *Tribology International*, vol. 84, pp. 90-99, 2015.

-
- [75] O. V. Thorsen and M. Dalva, "Failure identification and analysis for high-voltage induction motors in the petrochemical industry," *IEEE Transactions on Industry Applications*, vol. 35, no. 4, pp. 810-818, 1999.
- [76] R. Zhang, Z. Peng, L. Wu, B. Yao, and Y. Guan, "Fault Diagnosis from Raw Sensor Data Using Deep Neural Networks Considering Temporal Coherence," *Sensors*, vol. 17, no. 3, 2017.
- [77] M. Sohaib, C.-H. Kim, and J.-M. Kim, "A hybrid feature model and deep-learning-based bearing fault diagnosis," *Sensors*, vol. 17, no. 12, p. 2876, 2017.
- [78] S. Li, G. Liu, X. Tang, J. Lu, and J. Hu, "An ensemble deep convolutional neural network model with improved DS evidence fusion for bearing fault diagnosis," *Sensors*, vol. 17, no. 8, p. 1729, 2017.
- [79] V. Tra, J. Kim, S. A. Khan, and J.-M. Kim, "Incipient fault diagnosis in bearings under variable speed conditions using multiresolution analysis and a weighted committee machine," *The Journal of the Acoustical Society of America*, vol. 142, no. 1, pp. EL35-EL41, 2017.
- [80] S. Zhao, L. Liang, G. Xu, J. Wang, and W. Zhang, "Quantitative diagnosis of a spall-like fault of a rolling element bearing by empirical mode decomposition and the approximate entropy method," *Mechanical Systems and Signal Processing*, vol. 40, no. 1, pp. 154-177, 2013.
- [81] S. Yin, S. X. Ding, and D. Zhou, "Diagnosis and prognosis for complicated industrial systems—Part I," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 4, pp. 2501-2505, 2016.
- [82] S. Yin, S. X. Ding, and D. Zhou, "Diagnosis and prognosis for complicated industrial systems—Part II," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 5, pp. 3201-3204, 2016.
- [83] W. Huang, H. Sun, and W. Wang, "Resonance-based sparse signal decomposition and its application in mechanical fault diagnosis: A review," *Sensors*, vol. 17, no. 6, p. 1279, 2017.
- [84] X. Qin, Q. Li, X. Dong, and S. Lv, "The fault diagnosis of rolling bearing based on ensemble empirical mode decomposition and random forest," *Shock and Vibration*, vol. 2017, 2017.
- [85] M. Sohaib and J.-M. Kim, "A robust deep learning based fault diagnosis of rotary machine bearings," *Advanced Science Letters*, vol. 23, no. 12, pp. 12797-12801, 2017.

-
- [86] L. Frosini and E. Bassi, "Stator current and motor efficiency as indicators for different types of bearing faults in induction motors," *IEEE Transactions on Industrial electronics*, vol. 57, no. 1, pp. 244-251, 2010.
- [87] A. Soualhi, G. Clerc, and H. Razik, "Detection and diagnosis of faults in induction motor using an improved artificial ant clustering technique," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 9, pp. 4053-4062, 2013.
- [88] B. Eftekharijad, M. Carrasco, B. Charnley, and D. Mba, "The application of spectral kurtosis on acoustic emission and vibrations from a defective bearing," *Mechanical Systems and Signal Processing*, vol. 25, no. 1, pp. 266-284, 2011.
- [89] A. Widodo *et al.*, "Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine," *Expert systems with applications*, vol. 36, no. 3, pp. 7252-7261, 2009.
- [90] A. Glowacz, "Acoustic-Based Fault Diagnosis of Commutator Motor," *Electronics*, vol. 7, no. 11, p. 299, 2018.
- [91] P. Sangeetha and S. Hemamalini, "Dyadic wavelet transform-based acoustic signal analysis for torque prediction of a three-phase induction motor," *IET Signal Processing*, vol. 11, no. 5, pp. 604-612, 2017.
- [92] P. Nguyen, M. Kang, J.-M. Kim, B.-H. Ahn, J.-M. Ha, and B.-K. Choi, "Robust condition monitoring of rolling element bearings using de-noising and envelope analysis with signal decomposition techniques," *Expert Systems with Applications*, vol. 42, no. 22, pp. 9024-9032, 2015/12/01/ 2015.
- [93] G. Singh, T. C. Anil Kumar, and V. N. A. Naikan, "Induction motor inter turn fault detection using infrared thermographic analysis," *Infrared Physics & Technology*, vol. 77, pp. 277-282, 2016/07/01/ 2016.
- [94] M. Cerrada, R. V. Sánchez, D. Cabrera, G. Zurita, and C. Li, "Multi-Stage Feature Selection by Using Genetic Algorithms for Fault Diagnosis in Gearboxes Based on Vibration Signal," *Sensors*, vol. 15, no. 9, pp. 23903-23926, 2015.
- [95] H. Zhou *et al.*, "Weighted Kernel Entropy Component Analysis for Fault Diagnosis of Rolling Bearings," *Sensors*, vol. 17, no. 3, p. 625, 2017.
- [96] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, p. 436, 05/27/online 2015.
- [97] S. Wang, J. Xiang, Y. Zhong, and H. Tang, "A data indicator-based deep belief networks to detect multiple faults in axial piston pumps," *Mechanical Systems and Signal Processing*, vol. 112, pp. 154-170, 2018/11/01/ 2018.

- [98] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mechanical Systems and Signal Processing*, vol. 100, pp. 439-453, 2018/02/01/ 2018.
- [99] H. Shao, H. Jiang, H. Zhao, and F. Wang, "A novel deep autoencoder feature learning method for rotating machinery fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 95, pp. 187-204, 2017/10/01/ 2017.
- [100] Y. LeCun, "LeNet-5, convolutional neural networks," URL: <http://yann.lecun.com/exdb/lenet>, 2015.
- [101] P. Borghesani, P. Pennacchi, R. B. Randall, N. Sawalhi, and R. Ricci, "Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions," *Mechanical Systems and Signal Processing*, vol. 36, no. 2, pp. 370-384, 2013/04/01/ 2013.
- [102] R. Zhang, H. Tao, L. Wu, and Y. Guan, "Transfer Learning With Neural Networks for Bearing Fault Diagnosis in Changing Working Conditions," *IEEE Access*, vol. 5, pp. 14347-14357, 2017.
- [103] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.
- [104] J. Cai, S. Gu, and L. Zhang, "Learning a Deep Single Image Contrast Enhancer from Multi-Exposure Images," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2049-2062, 2018.
- [105] N. Yu, Z. Yu, F. Gu, T. Li, X. Tian, and Y. Pan, "Deep Learning in Genomic and Medical Image Data Analysis: Challenges and Approaches," *Journal of information processing systems*, vol. 13, no. 2, pp. 204-214, 2017.
- [106] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [107] X. Yu, E. Ding, C. Chen, X. Liu, and L. Li, "A Novel Characteristic Frequency Bands Extraction Method for Automatic Bearing Fault Diagnosis Based on Hilbert Huang Transform," *Sensors*, vol. 15, no. 11, pp. 27869-27893, 2015.
- [108] V. Tra, J. Kim, S. A. Khan, and J.-M. Kim, "Bearing fault diagnosis under variable speed using convolutional neural networks and the stochastic diagonal levenberg-marquardt algorithm," *Sensors*, vol. 17, no. 12, p. 2834, 2017.

-
- [109] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mechanical Systems and Signal Processing*, vol. 104, pp. 799-834, 2018/05/01/ 2018.
- [110] W. Ahmad, S. A. Khan, and J. Kim, "A Hybrid Prognostics Technique for Rolling Element Bearings Using Adaptive Predictive Models," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1577-1584, 2018.
- [111] R. K. Singleton, E. G. Strangas, and S. Aviyente, "The Use of Bearing Currents and Vibrations in Lifetime Estimation of Bearings," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1301-1309, 2017.
- [112] L. B. Cosme, W. M. Caminhas, M. F. S. V. D'Angelo, and R. M. Palhares, "A Novel Fault-Prognostic Approach Based on Interacting Multiple Model Filters and Fuzzy Systems," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 1, pp. 519-528, 2019.
- [113] H. T. Pham and B.-S. Yang, "Estimation and forecasting of machine health condition using ARMA/GARCH model," *Mechanical Systems and Signal Processing*, vol. 24, no. 2, pp. 546-558, 2010/02/01/ 2010.
- [114] L. Liao, "Discovering Prognostic Features Using Genetic Programming in Remaining Useful Life Prediction," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 5, pp. 2464-2472, 2014.
- [115] N. Li, Y. Lei, J. Lin, and S. X. Ding, "An Improved Exponential Model for Predicting Remaining Useful Life of Rolling Element Bearings," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 12, pp. 7762-7773, 2015.
- [116] B. Wang, Y. Lei, N. Li, and N. Li, "A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings," *IEEE Transactions on Reliability*, pp. 1-12, 2018.
- [117] J. Wu, C. Wu, S. Cao, S. W. Or, C. Deng, and X. Shao, "Degradation Data-Driven Time-To-Failure Prognostics Approach for Rolling Element Bearings in Electrical Machines," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 1, pp. 529-539, 2019.
- [118] A. Soualhi, H. Razik, G. Clerc, and D. D. Doan, "Prognosis of Bearing Failures Using Hidden Markov Models and the Adaptive Neuro-Fuzzy Inference System," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 6, pp. 2864-2874, 2014.
- [119] C. Chen, B. Zhang, G. Vachtsevanos, and M. Orchard, "Machine Condition Prediction Based on Adaptive Neuro-Fuzzy and High-Order Particle Filtering," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4353-4364, 2011.

- [120] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, "A Data-Driven Failure Prognostics Method Based on Mixture of Gaussians Hidden Markov Models," *IEEE Transactions on Reliability*, vol. 61, no. 2, pp. 491-503, 2012.
- [121] J. Ben Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, and F. Fnaiech, "Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network," *Mechanical Systems and Signal Processing*, vol. 56-57, pp. 150-172, 2015/05/01/ 2015.
- [122] J. A. K. Suykens, J. De Brabanter, L. Lukas, and J. Vandewalle, "Weighted least squares support vector machines: robustness and sparse approximation," *Neurocomputing*, vol. 48, no. 1, pp. 85-105, 2002/10/01/ 2002.
- [123] W. D. Mark and J. A. Hines, "Frequency-domain assessment of gear-tooth bending-fatigue damage-progression using the average-log-ratio, ALR, algorithm," *Mechanical Systems and Signal Processing*, vol. 45, no. 2, pp. 479-487, 2014.
- [124] T. H. Loutas, D. Roulias, and G. Georgoulas, "Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic e-support vectors regression," *IEEE Transactions on Reliability*, vol. 62, no. 4, pp. 821-832, 2013.
- [125] G. H. Golub and C. F. Van Loan, *Matrix computations*. JHU press, 2012.
- [126] G. Cauwenberghs and T. Poggio, "Incremental and decremental support vector machine learning," in *Advances in neural information processing systems*, 2001, pp. 409-415.
- [127] H. Q. J. Lee, G. Yu, J. Lin, and Rexnord Technical Services, ""Bearing Data Set"," U. o. C. IMS, Ed., ed. NASA Ames Prognostics Data Repository, <http://ti.arc.nasa.gov/project/prognostic-data-repository>: NASA Ames Research Center, Moffett Field, CA, 2007.
- [128] W. Ahmad, S. A. Khan, M. M. M. Islam, and J.-M. Kim, "A reliable technique for remaining useful life estimation of rolling element bearings using dynamic regression models," *Reliability Engineering & System Safety*, 2018/02/21/ 2018.
- [129] A. Saxena *et al.*, "Metrics for evaluating performance of prognostic techniques," in *2008 International Conference on Prognostics and Health Management*, 2008, pp. 1-17: IEEE.
- [130] X. Liu, P. Song, C. Yang, C. Hao, and W. Peng, "Prognostics and Health Management of Bearings Based on Logarithmic Linear Recursive Least-Squares and Recursive Maximum Likelihood Estimation," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1549-1558, 2018.

