DOCTOR OF PHILOSOPHY

DENOISING TECHNIQUES FOR VIBRATION SIGNAL ANALYSIS IN GEARBOX FAULT DIAGNOSIS SYSTEM
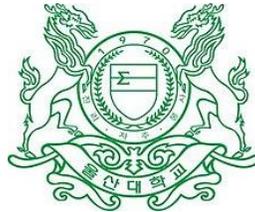
**The Graduate School
of the University of Ulsan
Department of Electrical, Electronic and
Computer Engineering**

NGUYEN CONG DAI

Denoising Techniques for Vibration Signal Analysis in Gearbox Fault Diagnosis System

Supervisor: Prof. Kim, Jong–Myon

A 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

NGUYEN CONG DAI

Department of Electrical, Electronic
and Computer Engineering,
University of Ulsan, Korea
February 2022

Denoising Techniques for Vibration Signal Analysis in Gearbox Fault Diagnosis System

This certifies that the dissertation of Nguyen Cong Dai is approved:

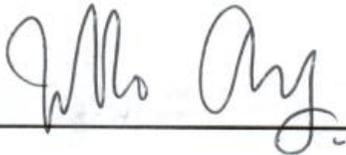_____

Prof. Kwon, Young-Keun,   Committee Chair

_____

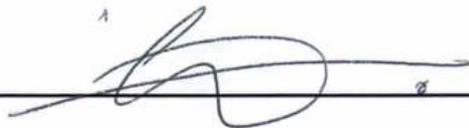Prof. Kim, Jong-Myon,   Advisor

_____

Prof. Jo, Dong-Sik,   Committee Member

_____

Prof. Chung, Jin-Ho,   Committee Member

_____

Prof. Kim, Cheol-Hong,   Committee Member

Department of Electrical, Electronic and Computer Engineering,
University of Ulsan, Korea
February, 2022

DECLARATION OF AUTHORSHIP

I, Nguyen Cong Dai, declare that this thesis titled " Denoising Techniques for Vibration Signal Analysis in Gearbox Fault Diagnosis System" and the work presented herein are my own. I confirm that:

- This work was done completely while in candidature for a research degree at the University of Ulsan, Republic of Korea.
- Where I have consulted the published work of others, it has always been clearly attributed to its original sources, and I have acknowledged all the sources.
- Where I have quoted from the work of others, the source is always mentioned. Except for such quotations, this thesis is entirely my work.
- Each research chapter of this dissertation relates to one SCI(E) indexed journal articles, which have been published.

Nguyen Cong Dai

UNIVERSITY OF ULSAN, REPUBLIC OF KOREA

February 2022

TERMS AND CONDITIONS RELATED TO COPYRIGHT

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g., Ph.D.) at the University of Ulsan, Republic of Korea. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

©2022 – Nguyen Cong Dai

VITA

Nguyen Cong Dai was born in Bac Ninh, Vietnam in 1982. He did his B.Sc. in Radio-Electronic Engineering from Le Quy Don Technical University, Hanoi, Vietnam in 2006, and M.Sc. in Electronic Control Systems from Sun Moon University, Republic of Korea in 2010. His interest in Masters was signal processing and image processing software and hardware co-design for combination systems. In 2006, he joined the Faculty of Radio-Electronic Engineering in Le Quy Don Technical University, Hanoi, Vietnam as a faculty member.

Since September 2018 he has been studied in the doctoral program in the faculty of computer engineering, Department of Electrical, Electronic and Computer Engineering, University of Ulsan, Republic of Korea under the supervision of Prof. Jong-Myon Kim. His current research interests contain vibration signal processing, artificial intelligence and deep neural network architectures for condition monitoring and fault diagnosis of gearbox systems and industrial machines.

I would like to dedicate this work to:

my family, with a special feeling of gratitude to my loving parents, my brothers, and my sisters, whose encouraging words push for tenacity ring in my study progression.

my children, for their joyfulness, childlike laughter's, loveable actions they provide to cheer up me during my doctorate program.

ACKNOWLEDGMENTS

During my Doctorate course at Ulan Industrial Artificial Intelligence Lab (UIAI), Department of Electrical, Electronic and Computer Engineering, University of Ulsan, Republic of Korea, I constantly achieved the wholehearted and heartfelt encouragement, attention and support from my supervisor, friends, and colleagues. Their favors have helped me to overcome the difficulties and challenges during my PhD course to get significant achievements.

First of foremost, I would like to express my profound gratitude to my advisor Professor Jong Myon Kim for his generous supports, constant guidance for my doctorate study and experimental researches, for his motivation, patience, and for his immeasurable sharing knowledge. During my doctorate period, he has given me the generous support both in technical and non-technical matters. I am beholden to my advisor for all his supervision, trust and kind support, and the financial assistance in the long journey of my PhD. Beside my advisor, I would like to gracefully thank to the members of my PhD supervisory committee for taking time to review my dissertation. I look forward to receive their valuable comments.

I would like to thank to my parents and my family members, who all have always loved and encouraged me in all situation of the life. Especially, I offer my appreciation and gratitude to my loving children for their efforts in the separation of a considerable distance during my PhD course.

I am grateful to my friends and my fellow lab members in the *Ulsan industrial Artificial Intelligence Lab (UIAI)*, University of Ulsan (UOU) for their cheerfulness, the technical discussions, and their enthusiastic support during my PhD.

Finally, I am thankful to the financial support of the BK21 plus program, the Korea Institute of Energy Technology Evaluation and Planning (KETEP), the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea, the ministry of Science, ICT and Future Planning, and the National Research Foundation of Korea (NRF), which enable me to pursue the doctorate course.

<div align="right">

Nguyen Cong Dai
University of Ulsan, Republic of Korea
February 2022

</div>

ABSTRACT

Gearbox fault diagnosis based on the vibration characteristic analysis has been widely developed and applied in research and industrial fields. Vibration characteristic of the gearbox systems accommodates the fault-related information of the rotating machines. In a simple gearbox, there are two engaged rigid blocks including a pinion wheel and a gear wheel for transferring the motion from the source to the load. The time variable gear-mesh stiffness causes the internal excitations during operation. If one or some of the teeth are faulty, abnormal movements will be appeared making the impulsive events in the vibration characteristic of a gearbox system. The vibration characteristic is normally sensed by vibration sensors (accelerometers) mounted on a gearbox sink. Thus the gearbox fault diagnosis model can be constructed based on the method of vibration signal analysis (VSA) to highlight the fault-related information. However, the gearbox vibration signals in real world are very complicated acting as non-stationary, nonlinear, and noisy overwhelming because of the inconsistent operation condition of a gearbox such as speed variation, fluctuation of the load, and the influence of mechanical resonances of other components. In addition, the several fault types create a similarly behavioral reflection on the gearbox vibration characteristic challenging the fault type discriminating process. Therefore, the enhanced denoising approaches and accurate fault type identifying methods are critically needed to construct the accurate and stable gearbox fault diagnosis model.

This study aims to propose novel adaptive denoising methodologies for filtering the original fault relative components from the raw noisy vibration signals regarding to variable operating speeds condition of a gearbox. Moreover, the AI based fault identification models are developed to process the fault-related information for accurately discriminating the fault types of a gearbox system. These advanced approaches of signal processing, feature engineering, and classification are incorporated to construct the sensitive and stable fault diagnosis frameworks for a multi-level gear fault gearbox under varying rotational speeds conditions. This thesis addresses four research keynotes.

For the first research topic, the adaptive noise reducer based Gaussian reference signal (ANR-GRS) approach is created for denoising the raw vibration signals. The ANR-GRS technique is established by following processes: first of all, the vibration characteristics of a gearbox system are structurally analyzed to classify noise types; then the speed-dependent Gaussian reference signals with adjustable parameters are generated, according to those noise types; finally, these generated noise-simulated reference signals were adaptively adjusted and accessed to the space between two consecutive fault-related frequency components and reduce the interference noise along with the whole frequency range of raw vibration signals. After denoising, the manual feature extraction method for extracting the optimized vibration subbands, outputted from the ANR-GRS, to many statistical features in time and frequency domains. Those fault representation features are used to input to a one-against-one support vector machine (OAOSVM) classifier for the fault type classification. The gearbox fault diagnosis scheme is validated for identifying the three fault types and one healthy state of the experimental testbed of a spur gearbox under varying speeds conditions. The result shows that the disturbance noises are significantly removed by using ANR-GRS method, thus feature extraction and OAOSVM based classifier provide excellent fault identification accuracy.

Secondly, the research topic focuses on the combined application of the adaptive noise control (ANC) method and genetic algorithm (GA) based feature selection to draw the sensitive fault diagnosis scheme. In this scheme, the applied adaptive noise control approach performs significantly removing noise elements and keeping original fault relative information from gearbox vibration signals. The outputs of ANC, optimized subbands, are then statistically extracted to many features configuring a feature pool. GA operates a heuristic searching process to select the most discriminative fault features, that represent samples of each fault type in clear separation allowing a simple machine learning model such as k-nearest neighbor (k-NN) for classifying defect categories into the respective types. This model is applied to classify six defective categories of a gearbox with multi-level gear defects. The accuracy result verifies the effectiveness of the combination fault identification model.

In the third research topic, the adaptive noise control (ANC) based Gaussian reference signal and stacked sparse autoencoder based deep neural network (SSA-DNN) are employed in combination for constructing a sensitive and speed invariant fault diagnosis model. The

applied model is used for diagnosing seven health states of the multi-level tooth cut gear defects (MTCG) gearbox under variable speed conditions. The deep learning model is built up by stacking the sparse autoencoder layers as the hidden layers and using a Softmax layer as the output layer of the network. SSA-DNN is capable to extract the spectra of the optimal vibration subbands, significantly denoised by ANC, into the high dimension feature pool of latent representative fault features, then selecting the most fault discriminant features for identifying the MTCG fault types under various speeds conditions. The effective evaluation of the proposed fault diagnosis scheme is verified by the classification result of the experiment on the vibration signal dataset of an MTCG gearbox collected under four different rotational speeds. The experiment is arranged by four sub-experiments using the datasets corresponding to four rotational speeds. In each sub-experiment, the network model is trained using a one-speed dataset and tested by two other speed datasets. The highest accuracy results are achieved, which outperform the state-of-the-art methodologies, validating the sensitive and speed invariance capabilities of the proposed fault diagnosis model in this research topic.

For the fourth study, the new localized adaptive denoising technique (LADT) is developed based on the ANR-GRS approach for improving the efficiency of noise reduction. Thus, an accurate and stable gearbox fault diagnosis scheme, that combines LADT with wavelet-based vibration imaging approach and deep convolution neural network model, is established. The new localized adaptive denoising technique results in optimized vibration subbands with negligible background noise. To obtain fault-related information, the wavelet-based vibration imaging approach (WVI) is applied to the denoised vibration signal. The wavelet-based vibration imaging approach decomposes the vibration signal into different time-frequency scales, these scales are reflected by a two-dimensional image called a scalogram. The scalograms obtained from the wavelet-based vibration imaging approach are provided as an input to the deep convolutional neural network architecture (DCNA) for discriminant features extraction and classification of multi-degree tooth faults (MDTF) in a gearbox under variable speed conditions. The proposed scheme outperforms the already existing state-of-the-art gearbox fault diagnosis methods with the highest classification accuracy.

Table of Contents

List of Figures

List of Tables

Nomenclature

2D	Two – Dimensional
AI	Artificial Intelligence
AC	Alternating Current
ANC	Adaptive Noise Control
ANN	Artificial Neural Network
ANR-GRS	Adaptive Noise Reducer based Gaussian Reference Signal
CBM	Condition-Based Monitoring
CM	Condition Monitoring
CNN	Convolution Neural Network
Cv	Convolutional layer
CWT	Continuous Wavelet Transform
DAQ	Data Acquisition system
DC	Direct Current
DCNA	Deep Convolution Network Architecture
DLM	Deep Learning Model
DNA	Deep Neural Architecture
DNN	Deep Neural Network
Do	Drop-out layer
DS	Drive Shaft
DT	Defect Type
DTA	Decision Tree Algorithm
DWT	Discrete Wavelet Transform
EMD	Empirical Mode Decomposition

H

FIR	Finite Impulse Response
Fc	Fully connected layer
FT	Fourier Transform
Fw	Framework
GA	Genetic Algorithm
GRS	Gaussian Reference Signal
ICA	independent component analysis
IIR	Infinite Impulse Response
IMF	Intrinsic Mode Function
HHT	Hilbert-Huang Transform
HT	Hilbert Transform
k-cv	k-fold cross validation
k-NN	k-Nearest Neighbor
LADT	Localized Adaptive Denoising Technique
LDA	Linear Discriminant Analysis
LMS	Least Mean Square
LPF	Low Pass Filter
MDFF	Most Distinguishable Fault Feature
MDTF	Multi-Degree Tooth Fault
MGTC	Multilevel Gear Tooth Cut
ML	Machine Learning
MTCG	Multi-level Tooth Cut Gear
MSE	Mean-Square Error
NDS	Non-Drive Shaft
OAOMCSVM	One-against-One Multi-Class Support Vector Machine

PCA	Principal Component Analysis
PFS	Principal Frequency Segment
Pm	Pooling layer
ReLU	Rectified Linear Unit
RPM	Revolution Per Minute
SFE	Statistical Feature Extraction
SSA-DNN	Stacked Sparse Autoencoder–based Deep Neural Network
STFT	Short Time Fourier Transform
SVM	Support Vector Machine
TBM	Time-Based Monitoring
TSA	Time Synchronous Averaging
t-SNE	t-Stochastic Neighbor Embedding
VSA	Vibration Signal Analysis
WA	Wavelet-based spectral Analysis
WBF	Window Bandpass Filter
WT	Wavelet Transform
WVI	Wavelet Vibration Imaging
WVT	Wigner-Ville transform

Chapter 1

Introduction

This chapter provides a brief background and fundamental explanation of conditional monitoring technique, vibration characteristics, related works, the motivations, and research objectives of this dissertation. All research work described in this thesis has been published in peer-reviewed conferences and journals. Moreover, in each chapter, introduction, background, and literature review, motivation for the given problem are also described. However, in this chapter, the background of a gearbox fault diagnosis system is explained in section 1.1, and the next sections (section 1.2, 1.3) described the related works, motivation, the outline of this thesis is presented in section 1.4.

1.1 Background

Gearboxes play a vital role in numerous application systems such as industrial machines, vehicles, robotics, electrical generators [1]–[3]. The gearboxes mainly function as increasing or reducing the rotational speed. However, gear failure can be prone to occur due to continuous and harsh operating conditions including fatigue phenomenon, exaggerated loads, inadequate lubrication, installation, and calibration problems [4], [5]. The failure of a gearbox can deteriorate the gearbox performance and entail damaging the rotation machine systems. These unwanted events might cause financial losses and human unsafety. Therefore, the appropriate maintenance processes are essential for fault early detection. These strategies can be performed by some methods as corrective and preventive maintenances [6]. Corrective maintenance is a very basic method as the implementation of repairing and modifying after machine faults occur, thus it can be applied for the small fault consequence. For the

applications with the large-scale and serious influence of risk and profit, the most suitable strategy is preventive maintenance. Preventive maintenance is used for preventing the defects of the rotation machinery and performed based on a time schedule (time-based maintenance - TBM) or based on condition (condition-based maintenance - CBM) [7]. However, TBM is done according to the calendar schedule, which is set for the maintenance process about historical information of the failure, regardless of the realistic health state of the rotation machine. Thus, TBM cannot prevent the defective effect on the machine when it is appeared out of the investigated periods. CBM is an effective maintenance technique to address the disadvantages of TBM by making maintenance decisions based on the current and actual health states of the items (e.g. gearbox, bearing, pump, …) in the rotation machines. Many researchers have demonstrated the advantages and efficiency of CBM in industrial applications [8]–[10].

The intrinsic failure of the items can be determined by conditional monitoring (CM) technique. Thus, CM is the key process of the CBM (the maintenance decision is more correct based on a more accurate CM process). CM can be understood as the activities, which can be manually or automatically established, designated to measure the parameters and characteristics, those represent actual health states of the items. Then CM is intended to detect the defects in the early phases for minimizing and avoiding the risk, secondary damages, and safety incidents. CM is applied in the gearbox fault diagnosis for early detection and avoiding the influence of defects in the gear teeth. Based on the application types of sensors and acquisition systems, different conditional monitoring techniques including vibration signal analysis, acoustics emission signal analysis, chemical monitoring, temperature signal analysis, electric current signal analysis are employed for the construction of the fault diagnosis systems. Nowadays, vibration monitoring is widely applied because of its easy arrangement and installation [11], [12].

In general, vibration monitoring is a non-destructive technique by using a vibration sensor (or accelerometer) mounted on the gearbox housing, for sensing the vibration characteristic, represented the actual states of a gearbox. Indeed, numerous health-state relative features of vibration characteristic of the item can be inspected in the vibration signals obtained from

sensors. The different fault states cause dissimilar patterns in the vibration characteristic. Therefore, analyzing the vibration signals allows us to discover the health conditions of the items of the machine during operation [13]. The aim of vibration monitoring for a gearbox is to detect and analyze any abnormal change in the vibration signal because of the gear failure condition to provide the premonition. Generally, a fault diagnosis system for the combination machine involves the recognition of fault elements or sub-elements of a whole system. Similarly, in a gearbox fault diagnosis system, the gear tooth is a vibrated dominant sub-element due to the gear mesh stiffness is the backbone of the gearbox vibration characteristic [14], [15]. If it is faulty (e.g. cracked tooth, worn tooth, spall), the operating properties of a gearbox are deteriorated causing catastrophic damages to other elements or the whole combination machine. Therefore, this thesis conducts the research on vibration monitoring for identifying gear tooth defects in several operating conditions. Figure 1.1 demonstrates the function block diagram of the general gearbox fault diagnosis systems based on the conditional monitoring approach.
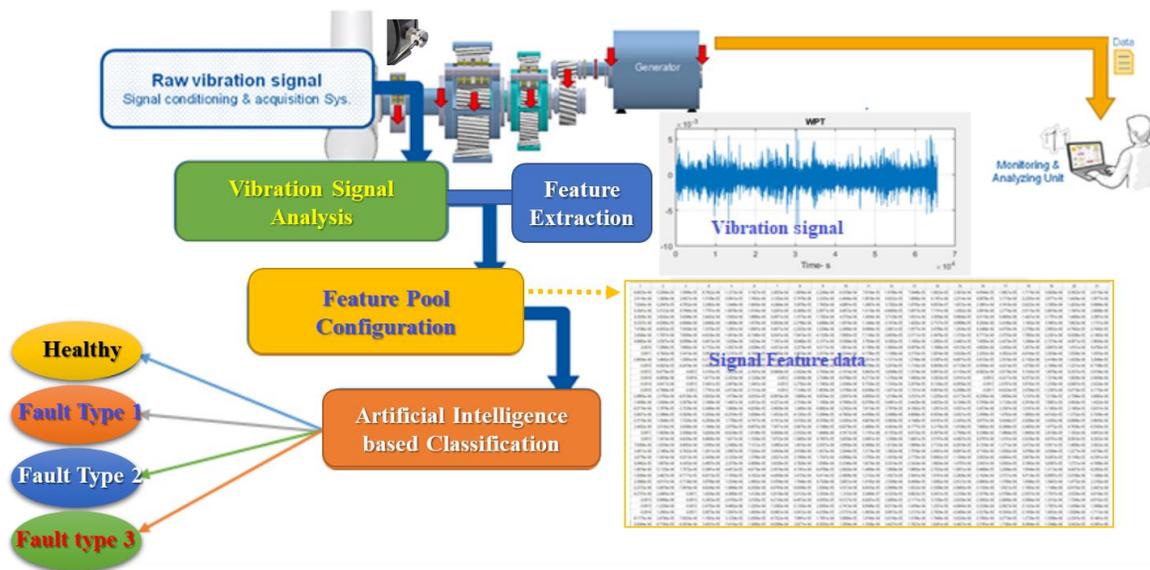
Figure 1.1. The general gearbox fault diagnosis system

As shown in Figure 1.1, it is constructed through four following main processes:

(1) Data collection: acquiring the vibration signal by using a vibration sensor and data acquisition system.

(2) Vibration signal analysis (VSA): this stage is to process the vibration signals and to capture the fault-related information.

(3) Feature pool configuration: the feature extraction and feature selection methods perform extracting and selecting the features of optimized vibration subbands obtained from the vibration signal analysis approach to configure feature pool.

(4) Fault pattern identification: the fault-related elements in the feature pool are used as input data of artificial intelligence based classification approaches for identifying the actual fault states.

The vibration characteristic of a gearbox in perfect and faulty cases have been analyzed in detail in [16]. In perfect condition, normal smooth and periodical movements of a pair gear generate the linear and periodic vibration signals, which are obtained from accelerometers and data acquisition systems. In a faulty case, the transferring motion is passed through a defective tooth causing impulsive events on vibration characteristics, thus the gearbox vibration signal can be a structure of a phase and amplitude-modulated signal [17]. This modulated signal comprises various frequency tones distributed around central frequencies in its spectrum. The series of central frequencies are meshing frequency harmonics, the sideband frequencies, and frequencies of free oscillations. These complex frequency components in the vibration signal can be used as fault relative information represented the actual health states of a gearbox [18], [19]. The aim of VSA, which is implemented by digital signal processing, is to highlight the informative ingredients from the vibration signals, which can be used for the identification of fault states of a gearbox system. Thus, VSA is the major process in the gear fault diagnosis application.

The gear faults are categorized into three major types: fabricating faults (e.g., incorrect tooth profile, eccentric calibrated wheels…), wrong setting and assembly, and operational faults caused by long time operation (e.g., broken tooth, cracked tooth, tooth spalling, worn tooth…). A conditional monitoring method is usually applied to diagnose operational fault types for detecting in an early phase. Thus, the experimental testbed of multi-level tooth cut gear (MTCG) faults gearbox under varying speed conditions was constructed in this study, shown

in Figure 1.2. The MTCG faults were created by cutting a tooth, which mounted on a gear wheel, in many depth levels (6.6%, 10%, 20%, 30%, 40%, and 50%) for simulating the same behavior of the gear failures caused by the long-term operation of a gearbox system. The gear wheel with 38 teeth is fixed with the load and non-drive shaft and engaged with a pinion wheel (25 teeth), which directly connects to AC motor through drive shaft, creating a gearbox with gear reducing ratio of 1:1.52. The rotation movement (torque) of the load is provided by the motion of the AC motor through the gearbox. Therefore, the rotational speed of the pinion wheel is equal to the rotational speed of AC motors, and the gear frequency is calculated by the pinion frequency and the gear ratio.
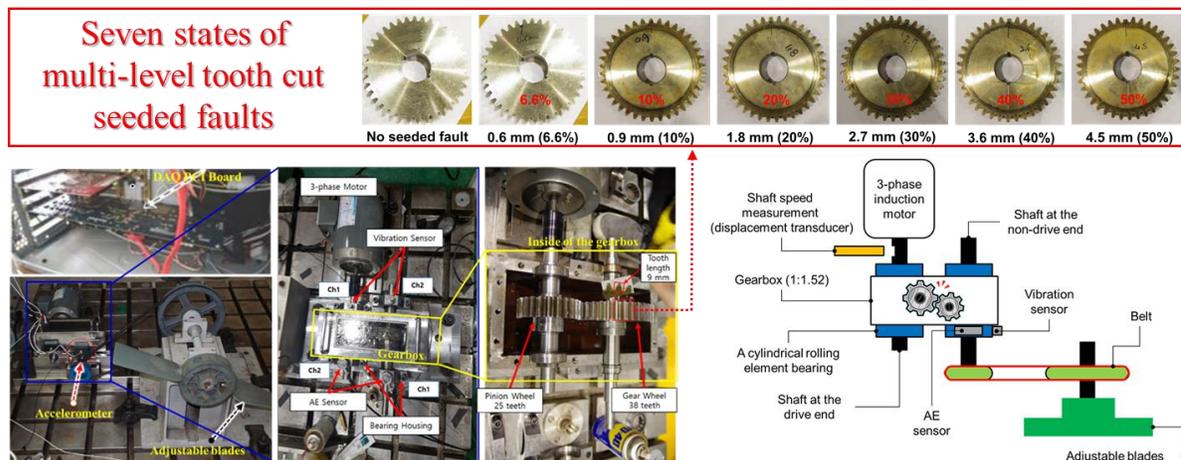
Figure 1.2. The arrangement of the experimental testbed

The datasets were collected by sampling vibration signal (using the accelerometer 622B01) with the frequency of 65536 Hz during one-second length. This process was repeated multiple times to acquire many vibration samples for each fault state (seven states) and alternately performed through four rotational speeds (300 RPM, 600 RPM, 900 RPM, and 1200 RPM).

Figure 1.3 demonstrates the analysis of a frequency spectrum of one vibration sample obtained from a gearbox with a gear tooth of 30% cut under the rotational speed of 900 RPM according to the gearbox vibration characteristic. The defect frequency tones are overwhelmed in noise components. Therefore, it is essential to develop the effective signal processing techniques, which are capable to highlight the informative components for identification of gear failures.

The defect frequencies:
- A meshing frequency (f_m) which can be computed by the parameter of pinion wheel (f_m = Number of pinion teeth × Rotational frequency of a pinion wheel) or using a gear wheel (f_m = Number of gear teeth × Rotational frequency of a gear wheel)
- A sideband frequency (f_g) is a frequency of a faulty wheel (if faulty tooth is mounted on a gear wheel, then it is termed as a sideband-gear frequency).

Figure 1.3. The analysis of the frequency spectrum of a vibration sample according to the vibration characteristic of a gearbox system

1.2 The previous researches

Vibration characteristic represents the response of the kinetic and dynamic models of a gearbox. In this matter, gear-mesh stiffness is the key factor for the exploration of fault gears [20], [21]. However, the gearbox vibration signals obtained from a vibration sensor in the combination systems are very complicated. Firstly, a gearbox usually works in the fluctuated or varying speeds condition according to the request of other systems, thus the gearbox vibration signals are considered as nonstationary and nonlinear signals [22]. On the other hand, the vibration characteristic of a gearbox is commonly impacted by undesired influences of the structural resonances of numerous lateral sub-systems such as gear wheels, shafts, bearings, and other mechanical elements, these induce the noise and redundant ingredients in the vibration signal making disorderliness of gearbox vibration characteristic. In addition, the interference noise from the sensor and data acquisition system is also affected to vibration signal during the data acquiring process [16], [23]–[25]. Those disturbance noises and unwanted residual elements, which are randomly occurred (i.e. random magnitude, random frequencies) in the whole range of vibration signals, significantly derange the gearbox vibration characteristic and defeat the clarifying process for fault useful components in the vibration signal analysis technique. In other words, the vibration signal is noisy, nonlinear, and non-stationary. Therefore, the applicable signal processing methodologies for removing redundancy and noise in vibration signals are previously necessitated.

In recent decades, numerous methodologies for analyzing the vibration signal have been proposed to find the fault representative elements in many domains. Those signal processing methods can be divided into several major types such as time domain VSA, frequency domain VSA, time-frequency domain VSA, and empirical mode decomposition (EMD) VSA, and other VSA.

For the time domain VSA, the performance is directly applied to the vibration signals in the time domain. The well-known method is to calculate the statistical features in the time domain as mean value, peak-to-peak value, kurtosis, skewness, root mean square, and crest factor. Then the evaluation process is used to examine the abnormal change which relates to the deterioration caused by the tooth faults. The description of statistical feature extraction in time domain is demonstrated in [14], [26], [27]. The other effective signal processing approaches in the time domain are time-synchronous averaging (TSA) and Hilbert transform. TSA is to extract the periodic portions by synchronous averaging the number of revolutions of the gears for detecting gear fault [28], [29], and Hilbert transform is carried to detect the repetitive impulsive amplitudes, formed by the operation of a faulty tooth, in the envelope [30], [31]. However, fault relative components are very difficult to be extracted from the noise immersed vibration signals by using the time domain VSA.

By contrast, the frequency domain VSA processes the vibration signal in frequency scope. Fourier transform (FT) is commonly used to transform the time domain vibration signal to a frequency spectrum. By using FT, the frequency tones are isolated to identify the fault relative elements. Many different techniques based Fourier transform have been developed in [32], [33]. There exist several frequency analysis methods as demodulated signal analysis and residual signal analysis. In a gear system, a vibration signal can be spectrally demodulated to figure out fault-related frequency tones (e.g., sideband frequencies, free oscillations) for discriminating the gearbox fault states. Demodulation analysis is developed in both the phase and magnitude of a vibration signal, which was recommended by McFadden in [34]. A detail explanation of demodulated analysis approach can be found in [35]–[37]. Another process technique in the frequency domain is residual signal analysis. The idea to develop this method is to remove the regular ingredients of a vibration signal, which are established in the normal operation condition, and to generate residual components for evaluation to detect fault impact.

There are some methods for obtaining the residual signals: removing meshing frequency harmonics [36], [38], [39], subtraction of two vibration signals acquired from the fault-free faulty gearboxes [26], [40], autoregressive model application [41], [42]. The frequency domain VSA can be efficient to distinguish a healthy gear and a fault gear, however, the frequency analysis is incapable to examine nonstationary signals.

To develop the appropriate signal processing methods for handling a nonstationary signal, time-frequency VSA has been conducted by many researchers. This method is carried out by applying the frequency transform technique on each divided timing window along the time length of a signal. Based on the method to create timing windows, different approaches have been constructed and widely applied including short-time Fourier transform (STFT), Wigner-Ville transform (WVT), and wavelet transform (WT) [27], [43]–[45]. Therein, WT is the most effective and well-known methodology for processing gearbox vibration signals. In WT, a vibration signal is decomposed to numerous vibration subbands by a family of wavelet functions, which are created by the transition and dilation process of the mother wavelet function (e.g., Morlet wavelet, Daubechies wavelet, Haar wavelet…). WT outcomes the good resolution analysis of the vibration signal in both time domain and frequency domain for the possibility of enhanced capturing short-term events of fault informative components [45], [46]. Nevertheless, the performance of WT might deteriorate in the case that informative frequency tones and mother wavelet function are uncorrelated, the disturbance noise tones, which are correlated with wavelet functions set, prevail to be excited rather than informative tones.

Another effective processing technique for nonstationary and nonlinear signals, which have been developed and applied for fault diagnosis systems of rotation machines, is a Hilbert-Huang transform [47]–[49]. This technique tries to decompose a vibration signal into set of intrinsic mode functions (IMFs), which represents fault-related intrinsic oscillations, by time adaptive performance of empirical mode decomposition (EMD). EMD has been demonstrated better performance than other time-frequency analysis methods as STHT and wavelet approach for VSA. Thus, the EMD technique was used in combination with others as: envelop analysis, wavelet approach to make use of their advantages [50]–[53]. However, the disadvantage of EMD is of noise sensitivity, thus in noisy circumstances, IMFs generated

from interference noise components cause a fault diagnosis system unreliably. The limitations have been assessed in [54], EMD has a better performance for low-noise vibration signals. Coupled with the aforementioned signal processing approaches, a denoising technique is of importance for analyzing vibration signals. There exist some noise filtering approaches as narrowband decomposition [55], discrete wavelet analysis [56], [57]. By applying these techniques, the interference noise in vibration signals has been reduced, but original fault useful frequency tones were also distorted due to attenuation of filters or threshold parameters.

The additional function module in the gear conditional monitoring system is fault pattern identification, which contains some functions as feature extraction, feature selection, and fault classification. Feature extraction is an interference operation that calculates the parameters of a vibration signal obtained from a vibration signal processing module. After extraction, the achieved parameters are configured as a feature pool. Next, the features in the pool are investigated and examined to figure out the characteristics of fault-related discrimination by the feature selection procedure. By selecting better representation features of gear health states, the dimension of the feature pool is reduced. The machine learning based classification approaches are applied for classifying the fault categories based on analysis of selected features. This is a process flow in a gearbox fault diagnosis model.

The conventional condition monitoring models have employed the manual feature extraction methods such as calculation of statistical parameters of vibration signal in time and frequency domain, complex envelope analysis, wavelet coefficient calculations… [58]–[60], and feature selection processes including principal component analysis (PCA), linear discriminant analysis (LDA), independent component analysis (ICA), genetic algorithm (GA) [61]–[65] for configuration of an optimal feature pool, then the optimal features are used as input data for fault classification by several shallow machine learning models: k-nearest neighbors (KNNs), decision tree, support vector machines (SVMs), and artificial intelligence networks (ANNs) [66]–[69]. The fault diagnosis models based on handcraft feature engineering and shallow networks have been of efficient performance for diagnosis of gear faults of a gearbox under the condition of constant rotational speed, however, those could restrict or degrade when processing the data of a gearbox under variable speeds conditions due to the limitation of feature pool configured by handcraft methods.

To address the issues of the dimensional limit of a feature pool, deep learning models (DLM) have been developed. DLM has demonstrated the advantages in many applications: computer vision, image processing, pattern recognition, and natural language processing… by the successfulness of discrimination of tiny different components for categories classification through numerous nonlinear transforming processes [70]–[73]. DLM performs automated feature extraction and feature learning procedures based on unsupervised hierarchy to configure the high dimension feature pools, which might contain discriminating features in small scales. Then, the fine-tuning process through neural layers of a DLM is executed to fetch the fault representation elements. Finally, gear fault types are identified based on those features. Further description of deep neural network architectures applied in a fault diagnosis system of a rotation machine can be found in [74]–[79].

1.3 Motivation

In order to construct the gearbox fault diagnosis system based on the conditional monitoring method, vibration signal analysis is particularly important. That is an initial step to obtain the gear state representative components, which highly influences the accuracy of the next steps of a fault identification system. Furthermore, a gearbox vibration signal is realized as a very complicated signal (i.e., nonlinear, nonstationary, and noisy), then the efficient VSA methods encounter the difficulties and challenges of the establishment. State-of-the-art methodologies, which have been described in the previous section (section 1.2), have been developed the VSA techniques for gearbox vibration signals, however, those studies were still associated with several issues that could depreciate the accuracy of the fault diagnosis system or could be inapplicable for a fault types identification system. The issues can be described as follows.

- There can be a problem to process nonstationary and nonlinear vibration signals.

- Noise reduction methods were not analyzed clearly in consideration of gearbox vibration characteristic

- There can be limitations for a gearbox operating under inconsistent working conditions, e.g. the rotational speeds of a gearbox were not considered.

- The informative frequency tones might be distorted and omitted.

- The challenge and inapplicability can be faced when processing vibration signals, which present similar behaviors of multi-level gear faults.

Therefore, the novel denoising methodologies have been constructed in this dissertation to analyze gearbox vibration signals for significant reduction of disturbance noise and preservation of fault informative elements. Those methods were constructed base on the adaptive denoising approaches as adaptive noise reducer-based Gaussian reference signal (ANR-GRS), and localized adaptive denoising technique (LADT).

Coupled with the signal processing approach, the feature engineering and classification procedures are also very essential to construct the conditional monitoring combination system for gearboxes. In this thesis, many techniques including a combination of handcraft feature pool configuration and shallow machine learning methods of classification, automatic and high dimensional feature engineering, and classification-based deep learning models, are experimentally employed to process de-noised vibration signals outputted from the proposed adaptive denoising technique modules.

The objective of this thesis is to develop effective noise controlling techniques and apply the proper neural network architectures to construct accurate and reliable fault identification systems for a multi-level tooth cut faults gearbox system under variable rotational speeds. This study is briefly described: (i) Construction and experimental verification of the novel effective denoising methodologies (ANR-GRS and LADT), (ii) Investigating the feature extraction, feature selection, and shallow neural network, (iii) examining and applying the deep neural network architectures for establishing a gearbox fault identification system.

1.4 Thesis Outline

The content of this dissertation is arranged by six chapters outlined below:

Chapter 1 briefly describes the background of a gearbox fault diagnosis system, conventional studies, motivation, and outline of this dissertation.

In **Chapter 2**, adaptive noise reducer based Gaussian reference signal methodology is developed for vibration signal analysis. A reliable gearbox fault diagnosis system is to establish by association of ANR-GRS, handcraft feature extraction, and support vector machine (SVM) approaches for identifying four health states of an MTCG gearbox under inconsistent speeds conditions.

Chapter 3 presents fault types identification and diagnosis model of a multi-level gear defect gearbox in the operation of a constant speed. In this model, the adaptive noise control technique is applied for denoising, genetic algorithm (GA) and k-NN are used for feature selection and fault classification, respectively.

Chapter 4 proposes a sensitive and invariant speed gearbox fault diagnosis model for seven fault types of an MTCG gearbox under varying speeds condition. In this model, adaptive denoising technique is used for significant removing noise and preserving fault useful information in the output optimized vibration subbands, then stacked sparse autoencoder based deep neural network (SSA-DNN) tries to perform automated feature extraction, fine tuning feature optimization and classification.

In **Chapter 5**, localized adaptive denoising technique (LADT) is developed based on ANR-GRS methodology for more efficient of noise reduction. An accurate and stable gearbox fault diagnosis framework is constructed based on LADT, wavelet-based feature pool configuration, and convolutional neural network architecture for identifying the fault states of a multi-degree tooth faults gearbox in the operation of the variable speeds.

Finally, the dissertation is of conclusion in **Chapter 6** with a summary of the contributions and a discussion of the future studies.

Chapter 2

A Reliable Fault Diagnosis Method for a Gearbox System with Varying Rotational Speeds

2.1 Introduction

Gearboxes are widely used in industrial applications, usually in harsh and continuous conditions, making them susceptible to a variety of failures. Defects can cause the gearbox system to break down and potentially damage complex mechatronic equipment or even cause a serious threat to safety, property, or customer satisfaction. Therefore, it is essential to diagnose gearbox faults regularly to ensure their early detection. The vibrations of gearbox systems have been studied since the 1980s, and previous researchers have found that gearbox vibrations have a keynote meshing frequency [3], [80] with complex sidebands around it and its harmonics [81], [82]. Therefore, the sideband frequencies and the meshing frequency and its harmonics are the informative components for identifying gear faults. Signal analysis is a backbone procedure for rotational-machine fault diagnosis research and applications. It works by decomposing the related fault features that are the groundwork for identifying fault patterns. The vibration characteristics of gearbox systems produce two major signals that can be analyzed for fault detection: acoustic signals and vibration signals [83]. Vibration signals are the most popularly used ones for gearbox fault monitoring because acquiring vibration data is easy [84]. However, vibration signals contain many types of noise from sources such as measurement systems (data acquisition systems), the environment, shafts, gears, and other related components and their impingement [85], [86]. All that noise, which exceeds the signal, fills the frequency spectrum of the vibration signal, and eclipses it.

Many signal processing methods using advanced techniques have been presented by many researchers: frequency analysis focusing on Fourier transform [87], Wigner distribution [88], rank-order morphological filter [89], cyclostationary signals for mechanical applications [90], and the envelope analysis [91], which is the most well-known for rotational-machine fault diagnosis applications such as bearing-fault diagnosis. It detects the repeating shock amplitudes that appear as faulty teeth traverse each rotation cycle. Using this method, the vibration signal is first processed by a bandpass filter to achieve a high signal-to-noise ratio, and second, the Hilbert transform is used to achieve the envelope. If the sideband frequencies in a gearbox vibration signal appear in the envelope, the presence of faulty teeth in the gearbox can be deduced [92], [93]. However, when the vibration signal is submerged in noise, it is difficult to recognize the informative components for fault diagnosis in the envelope.

Time-frequency analyses were developed to process non-stationary signals using a frequency transformation process divided based on windows across the time axis to capture informative events. The basic time-frequency analysis method is a short-time Fourier transform (STHT) or a spectrogram, such as a limited time window–width Fourier spectral analysis [94], [95]. The challenges of the STHT method, such as a failure of the assumption that the pieces of a non-stationary signal are stationary, difficulty adapting the observation window size to the size of a real stationary piece of signal, and the conflict between frequency resolution and time resolution (which is related to the Heisenberg uncertainty principle) limit its usability. To resolve the disadvantages of the STHT method, the wavelet approach was developed as an adjustable window frequency spectral analysis method. The basic wavelet function can be modified to meet special needs, so the wavelet transform produces outputs with good resolution in the low-frequency range and good time resolution in the high-frequency range [96]–[98]. In the region relevant for rotational-machine fault diagnosis, wavelet-based decomposition has been widely used to apprehend the useful components of a vibration signal in a non-stationary condition (in this context, non-stationary is the notion that the sideband frequency information of a vibration signal is time-variant). Wavelet transform decomposes a vibration signal into many sub-bands that express the time-frequency distribution through the dilation and transition of the mother wavelet. The sub-bands that contain fault-related intrinsic features can then be used in the fault diagnosis process [99], [100]. Nevertheless, the efficiency of the wavelet-based method correlates with the basic wavelet function, so

informative components that do not correlate as well with the applied wavelet could be missed or lost in the transformed outcome. In addition, the white noise that is frequently parasitic in a vibration signal and appears across the whole range of the frequency spectrum gives correlated oscillations with a high potential to appear as excitation. In this paper, the effect of noise on the wavelet applied to the wavelet transform method is compared with the proposed method for processing the signal path in the experimental results.

The Hilbert-Huang transform (HHT) was introduced as a better methodology for analyzing non-linear and non-stationary signals [47]. This technique is now often used for rotational-machine fault diagnosis [101]–[104]. The HHT method uses a time adaptive operation known as empirical mode decomposition (EMD) to decompose the signal into a group of complete and orthogonal components, denoted as intrinsic mode functions (IMFs), that represent the intrinsic oscillation modes of the fault-related components of a vibration signal. The HHT method was shown to outperform wavelet transform in rotational-machine fault diagnosis in [105]–[107]. To capture the advantages of the HHT, several fault detection tools combine EMD with other methods, such as envelope analysis and the wavelet-based technique. EMD and the envelope analysis combine in series: the vibration signal is first decomposed by EMD to determine the number of IMFs; the envelope analysis then processes the IMFs to monitor for fault-related components. Compared with previous methodologies, this combined technique had better results [105], [108], [109]. Combining wavelet transform and EMD for time-frequency analysis is another currently used combination method. It takes advantage of the strong points of the two techniques and minimizes their limitations, particularly aliasing in the high-frequency band (wavelet transform) and difficulties in isolating the signals within the second harmonic (EMD) [53], [110].

However, the EMD technique is sensitive to noise, so noise-related IMFs, which are not useful, can be generated by the EMD. As illustrated by Van M. et al. [54], EMD performs well in processing low-noise vibration signals and poorly in processing high-noise signals. In other words, even EMD combination techniques are unreliable in noisy environments. Therefore, to effectively apply enhanced signal analysis techniques to non-stationary vibration signals, a proper pre-processing method to reduce noise is required, such as narrowband demodulation [55] or discrete wavelet transform (DWT) [111], [112]. Applying those de-

noising methods effectively reduces the measurement noise, but the original informative signal is also distorted by the attenuation of a narrow bandpass filter in narrowband demodulation or the threshold in DWT-based de-noising. In other words, using one of those noise reduction methods can degrade the performance of a fault diagnosis system. Therefore, we have developed a new de-noising technique to reduce the noise from an original vibration signal by optimizing a process for filtering the weights and parameters of the reference input signals (adaptive) and considering the noise characteristics and rotational speed. We call our new technique the Adaptive Noise Reducer–based Gaussian Reference Signal (ANR-GRS).

The adaptive noise-controlling technique reduces noise by means of destructive interference. It consists of an adaptive filter and reference signals. The adaptive noise filter is a digital filter with an adaptive algorithm that adjusts the filtering coefficients (or tap weights) so the filter can be flexibly and optimally operated in unknown conditions with non-stationary signals to effectively remove low-level noise [113]. The typical performance criterion for adjusting the filtering weights (convergence condition) is based on the error signal, which is the difference between the output of the filter and the input reference signal as determined using the recursive least-squares or least mean square (LMS) algorithm. Between them, the LMS is more widely used because of its robustness and simplicity [114]. The ANR-GRS technique has three main function blocks: Gaussian reference signal (GRS) generation, adaptive noise filtering using the LMS algorithm, and optimal output sub-band selection. The generated GRS is a special signal consisting of a white-noise reference signal and a Gaussian reference signal with adjustable parameters (mean value and standard deviation) to identify noise components that are independent of the informative components in the frequency domains of the vibration signal from a varying speed gearbox. The adaptive noise filter consists of an M-tap digital Finite impulse response (FIR) filter and the LMS adaptive algorithm; it has two inputs: a reference input for the GRS signal with specific parameters and the desired input for a vibration signal. The noise-reduced sub-band is achieved as the output of the adaptive noise filter. The optimal output sub-band selection adjusts the parameter of the GRS signals to receive the set of noise-filtered sub-bands output by the adaptive filters and then selects the sub-band with the minimum mean square as the optimal sub-band, which is the final output of the ANR-GRS module. That output becomes the input for the feature pool configuration process used to extract the statistical features in the time and frequency domains of the

vibration signal as feature vectors [58] to be classified.

The heterogeneous feature pool improves the efficiency of gearbox fault expression for fault diagnosis process; however, the high dimensionality of the feature vectors can be a challenge for various machine learning techniques that can be used for decision making. In comparison with the other artificial intelligence algorithms, the classification performance of support vector machines (SVM) classifier is not much sensitive to the dimensionality of the feature vectors, in other words, this algorithm is not affected by the problem called 'curse of dimensionality'. Furthermore, SVM demonstrates excellent generalization performance, so this technique is capable of achieving high accuracy while classifying mechanical faults in rotation machinery [115]. Also, with an appropriate kernel function, SVM can accurately separate the non-linear datasets by hyperplanes in high-dimensional feature space using the non-linear mapping [116]. SVMs are widely used for fault diagnosis in many real-world applications [117]. They were originally designed for binary classification and then improved for multiclass classification using the one against one, one against all, or hierarchical strategy. Among them, the one against one strategy is the most reliable for our purposes [118]–[120]. Therefore, a one-against-one multiclass SVM (OAOMCSVM) is used in this proposed methodology.

The new hybrid technique employs the ANR-GRS, which produces an optimal sub-band, and then uses a machine-learning classification of fault types based on the OAOMCSVM on features extracted from that optimal subband to identify faults in a gearbox system. The experimental results show that the proposed method outperforms the aforementioned denoising methods, which verifies that the "clean" input can be used to produce correct output from the signal processing and classification paths.

The rest of this paper is organized as follows: Section 2 provides the characteristics of a gearbox vibration signal and the experimental test setup used in this study. The proposed methodology is explained in detail, from theory to the construction of the ANR-GRS, feature pool configuration, and OAOMCSVM classification, in Section 3. Section 4 demonstrates our experimental results in signal processing and classification. Finally, Section 5 concludes the paper.

2.2 The Characteristics of a Gearbox Vibration Signal and Experimental Testbed Setup

2.2.1 The Characteristics of a Gearbox Vibration Signal

The defects of a gearbox can be classified into three major categories: manufacturing defects (tooth profile error, eccentricity of the wheel, etc.), installation defects (parallelism), and operational defects (tooth wear, case wear, tooth spalling, tooth cracks). This research considers operational faults. A one-stage transmission gearbox, which consists of two rigid blocks with a pinion (on the drive side) and a gear (on the non-drive side), is illustrated in Figure 2.1. A healthy gear in normal condition working smoothly and periodically generates a linear and periodic vibration signal [81]. The vibration signal, s_h[mV/(m/s2)], of a fault-free normal pair of gears meshing under a constant load speed can be formulated as [121]:

$$s_h(t) = \sum_{i=0}^{N} S_i \cos(2\pi i f_M t + \varphi_i),\tag{2.1}$$

where S_i and φ_i are the amplitude and phase of the i-th meshing frequency harmonics; f_M is the meshing frequency (for the pinion: P is the number of teeth in the pinion wheel, f_P is the pinion rotation frequency, $f_M = P.f_P$; or $f_M = G.f_G$, G is the number of teeth in the gear wheel, f_G is the gear rotation frequency), and N is the total number of f_M harmonics in the frequency range of a vibration signal. Figure 2.2a shows the spectrum of the output vibration signal of a fault-free gearbox; it is filled with the frequency tones of the meshing frequency and its harmonics.
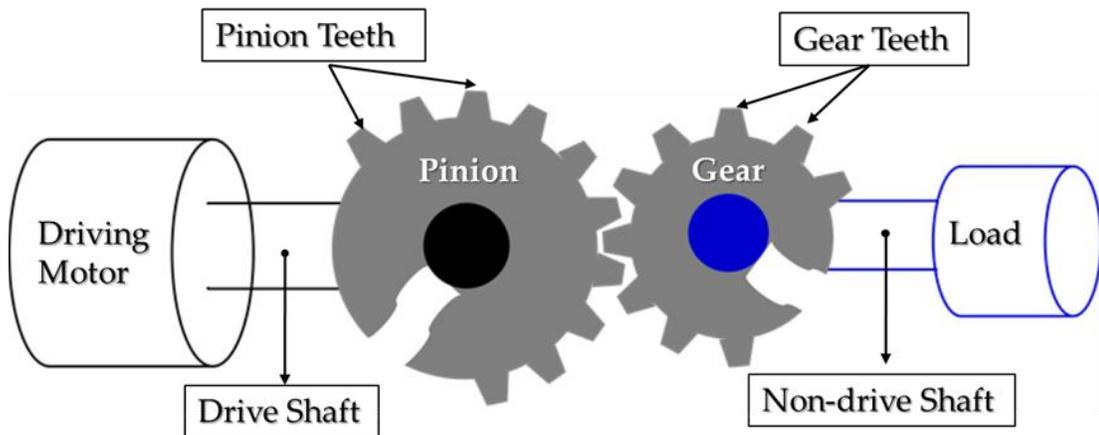
Figure 2.1. The spur gearbox model

In a fault case, when the motion transferred from the drive shaft to the non-drive shaft by the rotation between the pinion wheel and the gear wheel traverses a defective tooth (chipped, worn, or missing), an abnormal movement occurs that changes the impulses in the vibration signals. The vibration signal contains amplitude and phase modulations of the carrier frequency as the meshing frequency; its frequency spectrum includes sidebands, frequency components on two sides of the meshing frequency and its harmonics, as given in Equation (1). Thus, when the gear wheel has a faulty tooth, the velocity of the gear angle changes impulsively within the rotating functionality and generates a non-linear vibration signal in which issues such as speed variation, amplitude, and phase modulation prevail [82]. The vibration signal is formulated [81] as given by Equation (2.2), and an example of its spectrum is shown in Figure 2.2b:

$$s_f(t) = \sum_{k=0}^{N} S_k\big(1 + a_k(\mathrm{t})\big)\cos\big(2\pi k f_M \mathrm{t} + \varphi_k + p_k(\mathrm{t})\big). \qquad (2.2)$$

Here, $a_k(t) = \sum_{j=0}^{M} A_{kj}\cos(2\pi j f_G t + \mu_{kj})$ and $p_k(t) = \sum_{j=0}^{M} P_{kj}\cos(2\pi j f_G t + \xi_{kj})$. A_{kj}, P_{kj} are amplitudes and μ_{kj}, ξ_{kj} are phases of the j-th sideband in the amplitude and phase modulation signals, respectively, around k meshing harmonics.
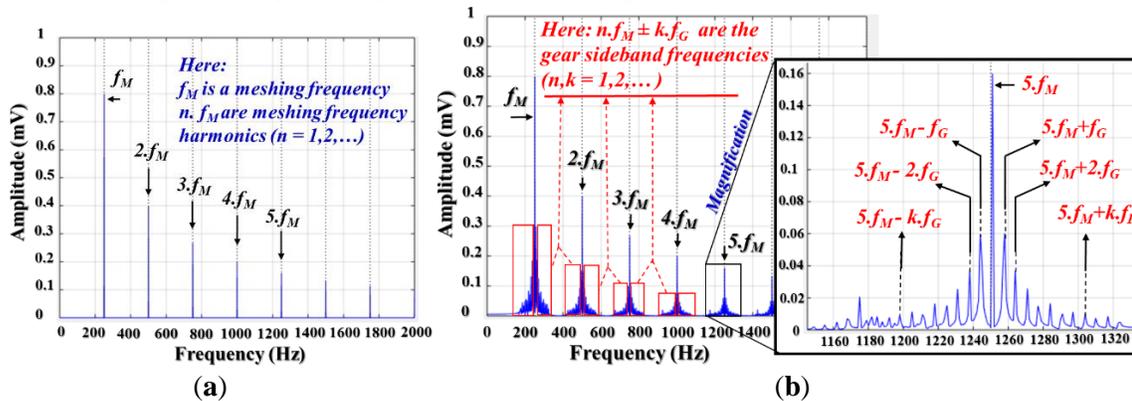
Figure 2.2. The frequency spectrum of the gearbox vibration signal: (a) a healthy gearbox and (b) a faulty gearbox

2.2.2 The Experimental Testbed Setup

The experimental testbed is illustrated in Figure 2.3. The pinion wheel is fixed to a three-phase AC induction motor by a drive shaft (DS). The motion (torque) is transmitted from the AC motor to the load as adjustable blades, which are mounted on the end of the non-drive shaft by the engaged teeth of a pinion wheel and a gear wheel (a gearbox with a gear reduction

ratio of 1:1.52).

The number of teeth on the pinion wheel is 25 (P = 25), the gear wheel has 38 (G = 38), and the length of each tooth is 9 mm. Figure 2.4 depicts the seeded tooth failures on the gear wheel: a perfect or healthy gear (H), tooth cut 10% (F1), tooth cut 30% (F2), and tooth cut 50% (F3). To measure the speed of the shaft rotation, a displacement transducer is placed to track the hole in the DS once per rotation. The vibration signals from the gear wheel in the normal condition and three levels of tooth cut defects (shown in Figure 2.4) were continuously acquired from the vibration sensor (accelerometer 622B01 made by the IMI Sensor Company) mounted on the end of the DS, 72.5 mm from the pinion gear. The analog vibration was digitized using PCI-2 data acquisition (the specifications of the data acquisition system are provided in Table 2.1). The sample datasets for each health condition (H, F1, F2, F3) of the gearbox under four shaft speeds are provided in Table 2.2.

(**a**)

(b)
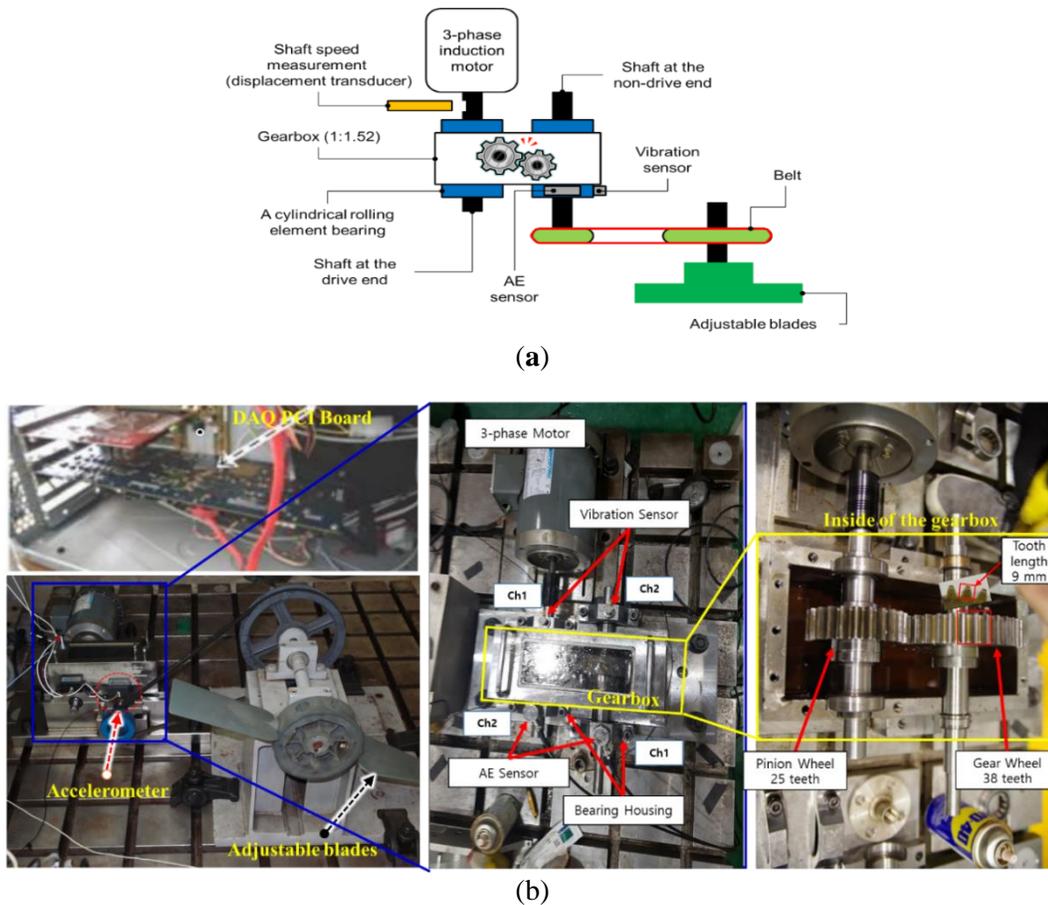
Figure 2.3. Experimental testbed setup: (a) function block diagram; (b) actual experimental assembly.
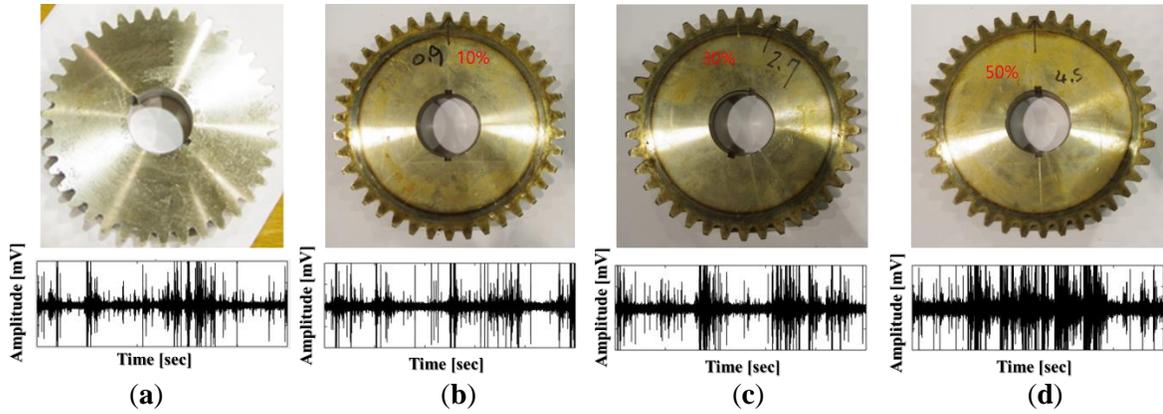
Figure 2.4. The health states of the gear wheel and examples of vibration signals at the rotation speed of 300 RPM: (**a**) no seeded fault, healthy gear, (**b**) tooth cut 10% (0.9 mm), (**c**) tooth cut 30% (2.7 mm), (**d**) tooth cut 50% (4.5 mm).

Table 2.1. Specification of the sensors and data acquisition system.

Device	Specification
Vibration sensor (Accelerometer 622B01)	Sensitivity (V/g): 10.2 mV/(m/s^2)
	Operational frequency range: 0.42 to 10 kHz
	Resonant frequency: 30 kHz
	Measurement range: ± 490 m/s^2
4- Channel DAQ PCI Board	18-bit 40MHz AD conversion, a sampling frequency of 65.536 kHz is used for each of two channels simultaneously
Displacement transducer	Distance from the head of a transducer to a hole: 1.0 mm
	Diameter of a hole: 12.80 mm
	Sensitivity: 0 to −3dB
	Frequency response: 0–10 kHz

Table 2.2. A detailed description of the fault types and dataset.

Gearbox health state	Description	Number of 1-s data samples acquired for each rotation speed				Sampling frequency (Hz)
		300 RPM	600 RPM	900 RPM	1200 RPM	
Healthy (H)	No seeded fault in the teeth of a gearbox	300	300	300	300	65536
Fault type 1 (F1)	Pinion tooth cut 10% (0.9 mm)	300	300	300	300	65536
Fault type 2 (F2)	Pinion tooth cut 30% (2.7 mm)	300	300	300	300	65536
Fault type 3 (F3)	Pinion tooth cut 50% (4.5 mm)	300	300	300	300	65536

Each health state is sampled by sampling frequency 65536 Hz in continuous 1 second (1-s sample) repeating by 300 times to receive 300 1-s samples for each shaft speed. Hence, the number of samples for each health state is 1200 vibration samples in four different shaft speeds, the total number of samples in this experimental testbed is 4800 of 1-s samples.

2.3 The Gearbox Fault Diagnosis Methodology

A function block diagram of the method for gearbox fault diagnosis proposed in this study is provided in Figure 2.5. We use four main processing blocks: the sensor and DAQ block, ANR-GRS, feature pool configuration, and multiclass SVM-based classification. To acquire discrete samples of each captured signal event containing information about the defective gearbox in the acquisition dataset, the raw vibration signal was sampled at a high frequency of 65536 Hz to acquire rich digitized vibration sample data under different shaft rotation speeds (300, 600, 900 and 1200 RPM), and the adjustable load was non-stationary. Though the operation frequency range of a vibration sensor in this study is from 0.42 Hz to 10 kHz (this is presented in Table 2.1), thus fault-related components in the frequency domain of vibration signals mostly exist in the lowest segment 0–10 kHz of their frequency spectrums. Therefore, the sampling frequency of the raw vibration signal (all vibration signals acquired in this paper) was reduced by a factor of three using a down-sampling technique. However, implementing decimation involves aliasing, so a low-pass Chebyshev Type I Finite Impulse Response filter (the filter with the order of 35 and a cut-off frequency of 10 kHz) was used for antialiasing [122]. The output sub-band signal from the lowpass filter (lpf(n), n is denoted as discrete-time) which have frequency spectrums in the range from 0–10 kHz, were then optimized by the ANR-GRS to achieve the optimal sub-band, opt(n), from which twenty-one features were extracted through a feature pool configuration, F(k), for classification by the OAOMCSVM.
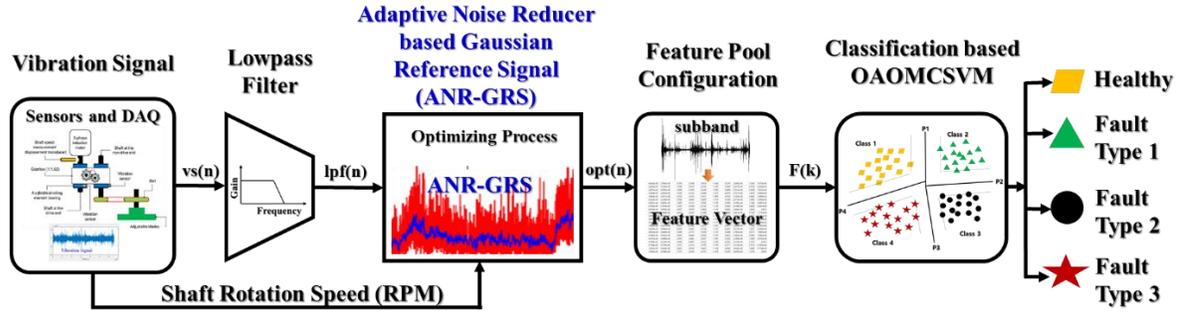
Figure 2.5. Function block diagram of the proposed methodology.

2.3.1 Adaptive Noise Reducer–based Gaussian Reference Signal

2.3.1.1. Adaptive Noise Filtering Technique

The digital filter

An adaptive filter combines the operation of a digital filter and an adaptive algorithm. The adaptive algorithm optimizes the coefficient (or weight) of a digital filter by using the feedback signal from the output (error signal) according to the signal condition or performance criteria [38]. Figure 2.6 illustrates the function of an adaptive filter constructed using a FIR filter and an adaptive algorithm. The output of the FIR filter is calculated as given in Equation (2.3):

$$g(n) \ = \sum_{m=1}^{M} c_m(n)r(n-m) = \ c^T(n)r(n) \qquad (2.3)$$

where, c_m, m = 0,1, …, M-1 (M is the digital filter length) are the adjustable weights (coefficients) of the filter, which do not depend on the sample time. The weight vector (M × 1) is formed as:

$$c(n) \equiv [c_0, c_1, …, c_{M-1}]^T, \qquad (2.4)$$

and r(n-m), m = 0, 1, …, M-1 are samples of an input signal composed of the vector M × 1:

$$r(n) \equiv [r(n), r(n-1), …, r(n-M+1)]^T. \qquad (2.5)$$
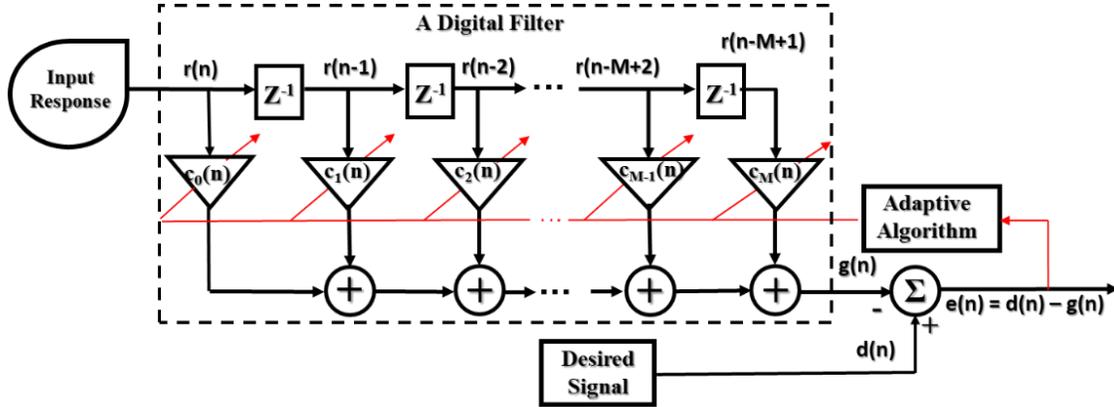
Figure 2.6. Function block diagram of an adaptive filter.

T denotes the transpose operation of the matrix. Then, the error signal e(n) is the difference between the FIR filter response, y(n), and desired signal, d(n), which can be calculated as:

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{g}(n) = \mathbf{d}(n) - \mathbf{c}^T(n)\mathbf{r}(n). \tag{2.6}$$

A common criterion for tuning the convergence of the weight vector, c(n), is the minimization of the mean-square error (MSE):

$$J \equiv E\{\mathbf{e}^2(n)\} = E\{[\mathbf{d}(n)-\mathbf{c}^T(n)\mathbf{r}(n)]^2\}$$

$$J = \mathbf{c}^T(n)R\mathbf{c}(n) - 2P\mathbf{c}^T(n) + E\{\mathbf{d}^2(n)\}, \tag{2.1}$$

where, $R \equiv E\{\mathbf{r}(n)\mathbf{r}^T(n)\}$ is the input autocorrelation matrix, and $P \equiv E\{\mathbf{r}(n)\mathbf{d}(n)\}$ is the cross-correlation vector between the input signal and the desired signal vector.

Equation (2.2) indicates that the MSE is a quadratic function of the filter weights (c), and its performance surface guarantees that it has a single global minimum MSE corresponding to the optimal vector $\mathbf{c}_o$. The optimal vector $\mathbf{c}_o$ can be found by taking the first derivative of Equation (2.3) and setting it to zero, the result achieved by Wiener-Hopf equation (assuming that R has an inverse matrix):

$$\mathbf{c}_o = R^{-1}P, \tag{2.8}$$

so that the minimum MSE is:

34

$$J_{min} = E\{\mathbf{d}^2(n)\} - P^T\mathbf{c_o}. \tag{2.9}$$

Adaptive algorithm

The adaptive algorithm is a recursive function to automatically adjust the coefficient vector, $c(n)$, to minimize MSE (J_{min}) so that the weight vector converges to the optimum solution, $\mathbf{c_o}$, after iteration loops. Both the LMS and recursive least-squares algorithms can be used to fetch the optimal solution [113], but the LMS is the most broadly used. To calculate the updated weight vector in the recursive loop, the LMS algorithm is based on the steepest-descent procedure using a negative gradient of the instant square error, which was devised by Widrow and Stearns [123] as follows:

$$c(n+1) = c(n) + \mu r(n)e(n), \tag{2.10}$$

where μ is the step size (or convergence factor) that determines the stability and convergence rate of the LMS algorithm. The algorithm adapts the weight vector to the optimal Wiener-Hopf solution ($\mathbf{c_o}$) given in Equation (2.9) by an iterative process with the convergence factor. The step size is selected in the range [114]:

$$0 < \mu < \frac{2}{MS_u}, \tag{2.11}$$

where S_u is the average power of the input signal $\mathbf{r}(n)$.

Adaptive noise filtering technique applied to a vibration signal

To construct the adaptive noise filter, the noise reference signal and observed signal are applied as the input signal of an adaptive filter (the input response in Figure 2.6) and the desired signal ($d(n)$ in Figure 2.6), respectively. The observed signal is the vibration signal acquired from the accelerometer sensor and digitized by the DAQ block reflecting gearbox behavior as expressed by the informative signals ($s(n)$) and the noise ($w(n)$), as shown in Figure 2.7. As explained in sections 1 and 2, the informative signals and noise are formed by different sources: the informative signal comes from the vibration of the gear and pinion teeth, whereas the noise comes from the measurement system, unrelated gearbox components, and mechanical resonances. Therefore, the informative signal and noise represent independent

processes ($E\{\mathbf{s}(n)\mathbf{w}(n)\} = 0$). To implement an effective adaptive noise filtering system, the generated noise reference signal, r(n), should meet two conditions (A and B):

    A) The generated noise reference, r(n), and informative signal, s(n), are uncorrelated and independent ($E\{r(n)s(n)\} = 0$).

    B) The characteristics of the generated noise reference, r(n), and noise, w(n), are homologous as much as possible.
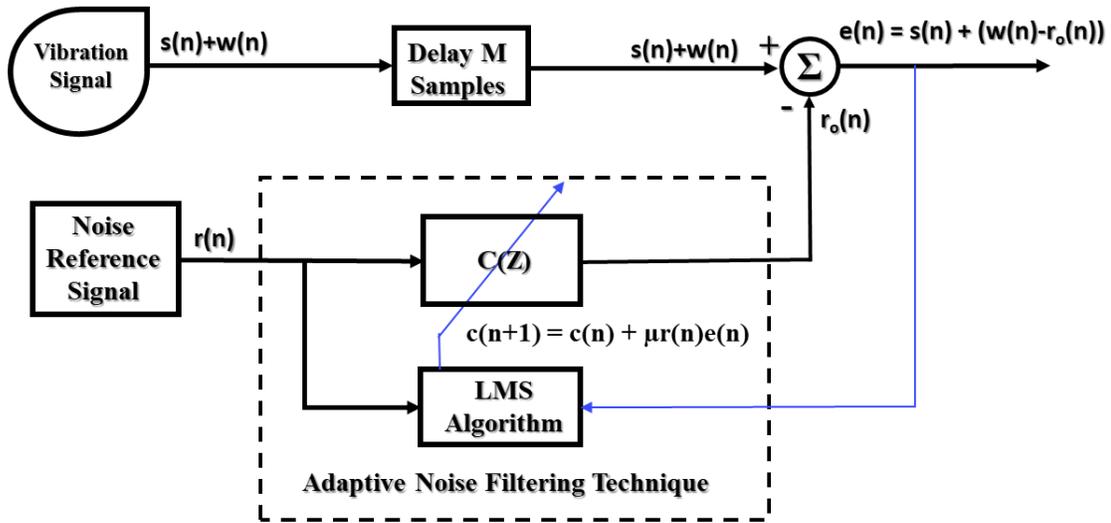
Figure 2.7. Function block diagram of an adaptive noise filtering technique

When those conditions are met, the MSE of the adaptive noise filter can be calculated as follows:

$$E\{\mathbf{e}^2(n)\} = E\{[\mathbf{s}^2(n) + (\mathbf{w}(n)\text{-}\mathbf{r}_o(n))]^2\}, \tag{2.4}$$

where $\mathbf{r}_o(n) = \mathbf{c}^T(n)\mathbf{r}(n)$, and the adaptive filter uses an FIR filter,

$$E\{\mathbf{e}^2(n)\} = E\{\mathbf{s}^2(n)\} + E\{[\mathbf{w}(n)\text{-} \mathbf{c}^T(n)\mathbf{r}(n)]^2\}. \tag{2.5}$$

The informative signal is independent of both the noise ($E\{\mathbf{s}(n)\mathbf{w}(n)\} = 0$) and the generated noise reference ($E\{\mathbf{r}(n)\mathbf{s}(n)\} = 0$). By implementing the LMS adaptive algorithm to adapt the filter coefficient vector, $\mathbf{c}(n)$, to the optimal vector, $\mathbf{c}_o$, the mean square of the output signal (error signal) approaches the single minimum of the performance surface. From Equation (2.6), the minimum MSE is taken to be:

$$\min_{\mathbf{c}(n)} E\{e^2(n)\} = E\{s^2(n)\} + \min_{\mathbf{c}(n)} E\{[w(n) - c^T(n)r(n)]^2\} \tag{2.14}$$

Therefore, the output signal of the adaptive noise filtering system carries the complete informative part of the gearbox vibration signal throughout the whole process of algorithm implementation. In addition, the noise integrated into the vibration signal is reduced; in the ideal case, the noise is removed ($\min_{\mathbf{c}(n)} E\{[w(n) - c^T(n)r(n)]^2\} = 0$). Therefore, for adaptive noise control, we implement the ANR-GRS.

2.3.1.2. ANR-GRS

In this paper, the noise ($w(n)$) in the gearbox vibration signal is divided into two types: white noise ($u(n)$) and band noise ($b(n)$). The white noise arises from the measurement system: the amplifier, detector, DC power supply, thermal vibration of the semiconductor atoms, etc. In the frequency domain, the power of the white noise is spread across the whole frequency spectrum of the vibration signal (theoretically, the power of white noise is spread from $-\infty$ to ∞ in the frequency axis) [86]. Band noise, on the other hand, represents noise caused by unrelated components [85]. The frequency harmonics of the band noise are distributed around the informative components of the gear sideband frequency, meshing frequency, and their harmonics. Therefore, the informative signal inside the vibration signal is separately independent of both types of noise. The ANR-GRS module is built using the adaptive noise filtering technique and reference noise-related generation signals, as illustrated in Figure 2.8. To reduce the white noise, we apply a generated white noise signal with a uniform, random distribution function ($v(n)$). The oscillation form of the generated white noise is thus analogous to the white noise integrated into the vibration signal. Because its frequency spectrum is within the observed frequency range, the maximum level of the power spectrum average (PSA) of the reference white noise is reduced to less than 10% (10% in this study) of the PSA of the vibration signal to ensure that the informative signal can be eligible for

37

conditions A or B. The GRS (g(n)) is created to adapt to the band noise inside the vibration signal. To make the proposed methodology as an invariant model, the GRS generation module uses the shaft rotation speed (RPM) information from the displacement transducer and the vibration signal as the input parameters. Then, the mean frequency (F_{Center}) and the standard deviation of the GRS are calculated based on the frequency of the defective wheel, which is a function of RPM (the gear frequency in this paper). The GRS window is confined entirely within the frequency space between two consecutive sideband frequencies (a sideband segment), pictorially described in Figure 2.9, and computed as follows:

$$W_{GRS}(k) = \sum_{k=1}^{N_b} e^{-\frac{(k-F_{Center})^2}{2\Delta}},$$ (2.15)

where $\Delta = \sigma^2$ is the variance, σ is the standard deviation of the GRS window, and F_{Center} is the mean frequency of the GRS window. They function as the frequency of a faulty wheel (the gear frequency, $f_G = P.RPM/G$ in this research).

$$F_{Center} = \alpha. f_{FW}.$$ (2.16)

By linearization of the Gaussian function, the standard deviation (the characteristic of a Gaussian distribution) can be approximately calculated as:

$$\sigma = 0.318. F_{Center} = 0.318. \alpha. f_{FW}.$$ (2.17)

N_b is the number of frequency bins in a sideband segment and defined as follows:

$$N_b = \frac{2N_s}{F_s}. f_{FW},$$ (2.18)

where N_s is the number of samples of the vibration signal, F_s is the sampling frequency of the vibration signal, and f_{FW} is the frequency of the faulty wheel (the gear frequency, f_G, in this paper).
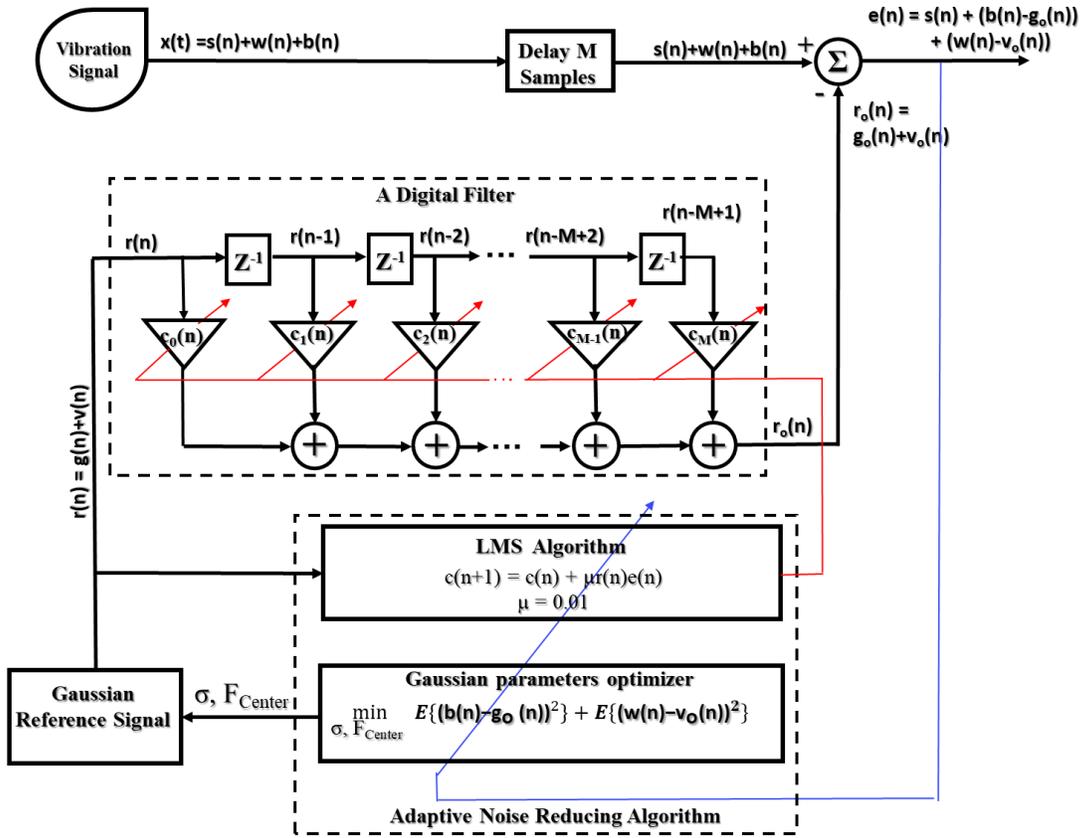
Figure 2.8. A function block diagram of the ANR-GRS module and parameter adjustments
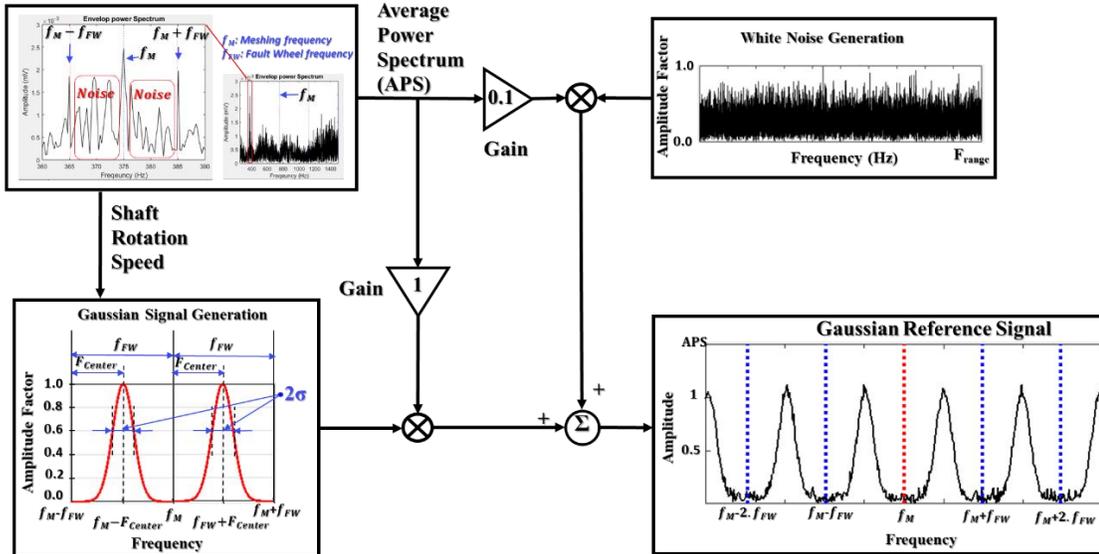
Figure 2.9. The overall flow chart of GRS signal generation for the ANR-GRS module

To qualify condition B, the frequency components of a Gaussian window are separated from the informative frequencies (sideband frequencies). A Gaussian window is placed completely inside the space between two continual sideband frequencies in the visualization. Thus, the adaptation process for a band-noise reduction is to preserve the original informative frequency component (significantly reducing the noise components and causing negligible attenuation of the informative components). From Equations (2.15), (2.16), and (2.17) and Figure 2.9, the coefficient α is selected in the range from 0.25 to 0.75, and the qualified Gaussian window signals are generated using the parameters in the following ranges:

The range of the mean value:

$$0.25. \, f_{FW} \leq \ F_{Center} \leq 0.75. \, f_{FW}. \tag{2.19}$$

The range of the standard deviations of the Gaussian reference signal:

$$\sigma = \begin{cases} 0.318 \,.\, \alpha. \, f_{FW} & when \; 0.25 \leq \alpha \leq 0.5 \\ 0.318 \,.\, (1-\alpha). \, f_{FW} & when \; 0.5 < \alpha \leq 0.75 \end{cases} \tag{2.20}$$

Therefore, the implementation of a stepping adjustment in the coefficient α drives a change in the mean value and standard deviation (the position and shape) of the Gaussian window, which defines the condition for fetching the optimal Gaussian window, as illustrated in Figure 2.9.

2.3.1.3 The process for calculating the optimized subband

First, the ANR-GRS algorithm germinates the initial parameters for the Gaussian signal generation module: starting value of α=0.25 in this paper, adaptive filter (M-tap, M=40 in this study), coefficient vector c(n) = [0,0,..,0], and step size μ (μ=0.01). The parameter α is scanned in the range [0.25 0.75] in steps of 0.01 in company with the input rotation speed (RPM) to compute the F_{Center} (mean value) and standard deviation (σ) using Equations. (2.16) and (2.17). To generate the specific GRS needed for the reference input of the adaptive filter r(n), the output of the adaptive filter is connected to the minus port of the summation module, r_o(n). The vibration signal, which contains both the informative component and noise, is entered as the desired input and delayed for M sampling time steps to be compatible with the delayed

processing of the FIR digital filter. The LMS algorithm adjusts the coefficient vector to receive the LMS of the error, which is the output of the summation module. The output error signal, which has LMS (and to which the optimal coefficient vector is set), is pushed into the set of proposed optimized sub-bands.

Finally, the algorithm calculates the mean square value of each subband in that set and then selects the sub-band with the minimum value as the optimized subband and output of the AND-GRS module (Figure 2.10).
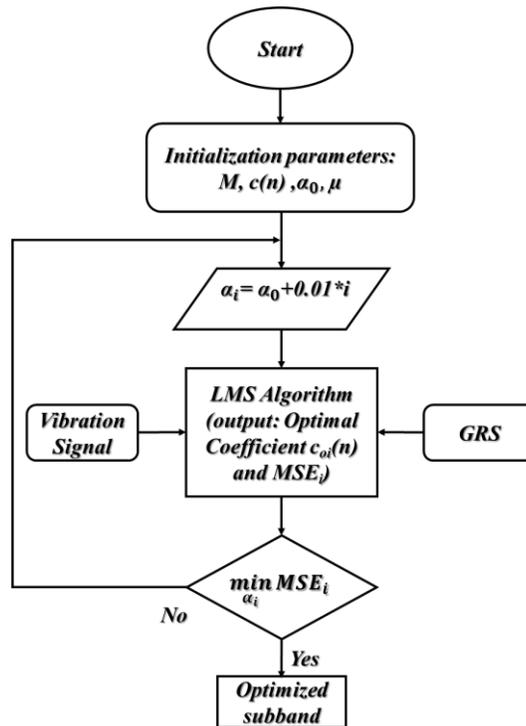
Figure 2.10. The algorithm flow chart of the ANR-GRS module.

2.3.2. Feature pool configuration

We found the ANR-GRS methodology to be highly effective in reducing most of the noise components from a 1-s raw vibration signal while leaving the information about gearbox faults intact. The optimized subband output from the ANR-GRS, i.e., the "clean" signal presenting the characteristics of the gearbox component vibration with trivial noise effects, carries the intrinsic fault symptoms of the cut tooth defects. We then use those optimal sub-bands, rather than the raw 1-s vibration signals, to extract features. According to Caesarendra et al. [124],

the statistical parameters from the time and frequency domains of the signal are congruent and subservient for fault classification using machine learning. Table 2.3 displays twenty-one features, eighteen time-domain features (e.g., root means square, square mean root, kurtosis, skewness, margin, impulse, and peak to-peak value) and three frequency-domain features (root mean square frequency, frequency center, and root variance frequency) for each optimal subband. The feature pool dimensionality is $N_{HS} \times N_{1\text{-SEC}} \times N_F$, where N_{HS} is the number of gearbox health states (number of classes) that need to be classified (4 classes in this study: healthy, pinion tooth cut 10%, pinion tooth cut 30%, and pinion tooth cut 50%), $N_{1\text{-SEC}}$ is the number of 1-s samples of each class (300 in this study), and N_F defines the number of features (21 in this study). Therefore, groups of 21-feature vectors were considered as the validating input dataset for our proposed intelligent fault-detection method based on a multiclass SVM.

Table 2.3. Definition of statistical features in the time and frequency domains

Features	Equations	Features	Equations	Features	Equations
Peak	$Max(\lvert s \rvert)$	Shape factor	$\dfrac{s_{rms}}{\frac{1}{N}\sum_{n=1}^{N}\lvert s_n \rvert}$	Mean $(\bar{s})$	$\dfrac{1}{N}\sum_{n=1}^{N} s_n$
Root mean square(s_{rms})	$\sqrt{\dfrac{1}{N}\sum_{n=1}^{N} s_n^2}$	Entropy	$-\sum_{n=1}^{N} p_n \cdot log_2(p_n)$	Shape factor square mean root	$\dfrac{s_{srm}}{\frac{1}{N}\sum_{n=1}^{N}\lvert s_n \rvert}$
Kurtosis	$\dfrac{1}{N}\sum_{n=1}^{N}(\dfrac{s_n - }{\sigma}$	Skewness	$\dfrac{1}{N}\sum_{n=1}^{N}(\dfrac{s_n - \bar{s}}{\sigma})^3$	Margin factor	$\dfrac{max(s)}{s_{smr}}$
Crest factor	$\dfrac{Max(\lvert s \rvert)}{s_{rms}}$	Square mean root (s_{smr})	$(\dfrac{1}{N}\sum_{n=1}^{N}\sqrt{\lvert s_n \rvert})^2$	Peak to peak	$max(s)-min(s)$
Clearance factor	$\dfrac{Max(\lvert s \rvert)}{s_{smr}}$	5th normalized moment	$\dfrac{1}{N}\sum_{n=1}^{N}(\dfrac{s_n - \bar{s}}{\sigma})^5$	Kurtosis factor	$\dfrac{Kurtoris}{s_{rms}^4}$
Impulse factor	$\dfrac{Max(\lvert s \rvert)}{\frac{1}{N}\sum_{n=1}^{N}\lvert s_n \rvert}$	6th normalized moment	$\dfrac{1}{N}\sum_{n=1}^{N}(\dfrac{s_n - \bar{s}}{\sigma})^6$	Energy of signal	$\sum_{n=1}^{N} s_n^2$
Frequency center (FC)	$\dfrac{1}{N_f}\sum_{f}^{N_f} S(f)$	Root mean square frequency	$\sqrt{\dfrac{1}{N_f}\sum_{f}^{N_f} S(f)^2}$	Root variance frequency	$\sqrt{\dfrac{1}{N_f}\sum_{f}^{N_f}(S(f) - FC)}$

where s is an input signal (i.e., optimized subband), N is the total number of samples, S(f) is the magnitude response of the fast Fourier transform of the input signal s, N_f is total number of frequency bins, $\sigma =$

$$\sqrt{\frac{1}{N}\sum_{n=1}^{N}(s_n - \overline{s})^2}, \ and \ \ p_n = \ \frac{s_n^2}{\sum_{n=1}^{N} s_n^2}$$

2.3.3 Gearbox fault classification using a multiclass SVM classifier

The principle operation of an SVM is based on the statistical learning theory of Vapnik [118] and quadratic programming [119]. It was actually designed to classify binary datasets by finding the optimal plane, generally called the hyperplane, with the largest margin-gap separating it from both binary classes.

Let $\{(x_m, y_m), m=1, 2, ..., M\}$ be the given training dataset with M samples, where each sample data $x_m, \in \mathbb{R}^D$, $\mathbb{R}^D$ is a D-dimensional feature vector, and y_m ($y_m \in \{-1, +1\}$) are the class labels. The SVM is used to find a set of linearly separable hyperplanes between two classes and maintain the maximum distance (called the margin) from both of them. The hyperplane, denoted as w, is determined as the maximized width of the margin and the minimized structural risk, given by

$$(w, b) = \underset{w,b}{\text{argmin}} \frac{1}{2} w^T w + C \sum_{m=1}^{M} \xi_m, \tag{2.21}$$

subject to: $\quad y_m(w^T \psi(x_m) + b) \geq 1 - \xi_m, \forall m = 1,2,.., M;$

$$- \xi_m \leq 0, \forall m = 1,2,.., M$$

Here, b is bias, C is the trade-off parameter, $\xi = \{\xi_1, \xi_2, ..., \xi_N\}$ is the set of slack variables, and $\psi(.)$ is a feature vector in the expanded feature space. Equation (2.21) can be solved by applying the Lagrange duality solution [117] as shown below:

$$\underset{\alpha}{\text{argmax}} \ w(\alpha) = \sum_{m=1}^{M} \alpha_m - \frac{1}{2} \sum_{m=1}^{M} \sum_{k=1}^{M} \alpha_m \alpha_k y_m y_k \psi^T(x_m) \psi(x_k) \tag{2.22}$$

subject to: $\sum_{m=1}^{M} \alpha_m y_m = 0, 0 \leq b_m \leq C, \ \forall m = 1,2,..., M$

where, α_m and α_k are Lagrange multipliers, x_m and x_k are two input training vectors, and $K(x_m, x_k) = \psi^T(x_m)\psi(x_k)$ is a kernel function used to map the input data space into a higher-dimensional feature space. Several kernel functions, such as linear, polynomial, Gaussian, radial basis, and sigmoid functions, can be used in SVM classification methods. Countless

classification applications have more than two classes in their datasets and thus require a solution beyond the binary SVM just described. Multiclass SVMs have been developed to classify datasets of N different classes (N>2), and they use one of three structures: one-against-one, one-against-all, and hierarchical. Among those structures, OAOMCSVM requires more classifiers than the others, but it also has the most reliable classification accuracy [118]. Therefore, we use OAOMCSVM, illustrated in Figure 2.11, in the methodology proposed in this paper.
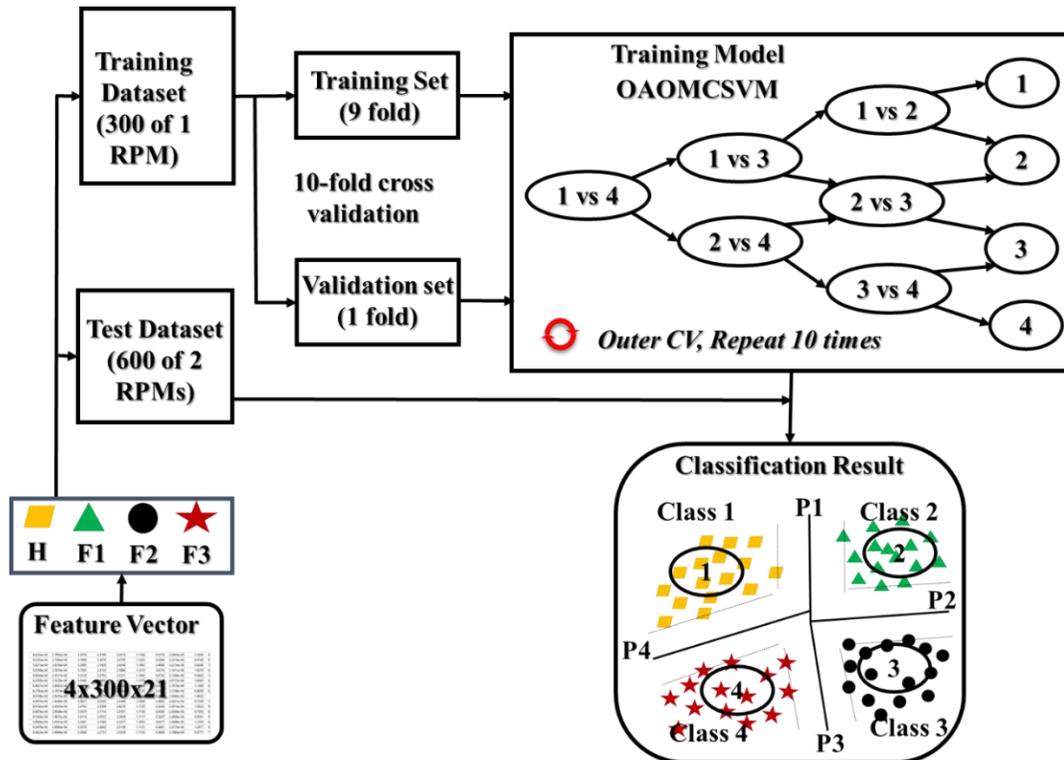
Figure 2.11. The classification methodology of OAOMCSVM

2.4 Experimental Results

To verify the advantage of the ANR-GRS module in the proposed methodology, we implemented experiments in two technological zones, signal processing and features dataset classification, and compared our results with those from conventional methodologies.

2.4.1 Signal Processing Experimental Results

The 1-s vibration signals acquired using the experimental testbed described above contained the phase and amplitude modulation signals bearing information about the health states of a gearbox. To investigate the effectiveness of the proposed noise reduction technique, the experimental dataset was collected under various shaft rotation speeds that are equal to 300, 600, 900, and 1200 RPM, respectively. The vibration signals output from the accelerometer are analog, so they were digitized with a high holding sample frequency of 65536 Hz to gather as much information (and noise) as possible on the wideband PCI-based data acquisition board (Table 1). Each 1-s digital vibration signal was down-sampled three times, incorporating a low-pass filter for antialiasing to output a digital vibration signal realistically compatible with the working range frequency of the accelerometer (0–10 kHz, Table 1). Then, the vibration signal was input into the ANR-GRS module (Figure 2.5). The shaft rotation speed (RPM), measured by the displacement transducer, was observed by the ANR-GRS module according to appropriate vibration signal data to generate the Gaussian reference signal. The optimal subbands were the output of the ANR-GRS module.

To demonstrate the superiority of the ANR-GRS technique, we compared its optimized subbands with the outputs of other signal processing approaches for noise reduction: The Hilbert transform (HT), window bandpass filter (WBF), and wavelet transform with optimal subband-based maximum kurtosis (WTK). We tested those approaches by replacing the ANR-GRS module with them. Figure 2.12 illustrates the frequency spectra compared with the input vibration signal. Figure 2.12 shows the output of a low-pass filter that received a 1-s vibration sample with 900RPM (15Hz) of fault type 2 (meshing frequency, f_M = P.RPM = 25.15 = 375 Hz and sideband gear frequency, f_G=P.RPM/G=9.87 Hz, shown as lpf(n) in Figure 2.5 and labeled as the OutLPF signal in Figure 2.12a). The output signals from the noise-reduction modules are shown in Figure 2.12b (OutHT signal), Figure 2.12c (OutWBF signal), Figure 2.12d (OutWTK signal), and Figure 2.13, the proposed ANR-GRS (OptANR signal). The three conventional methods (HT, WBF, WTK) changed the outLPF signal into different shapes and types (the outLPF signal is an amplitude and phase modulation signal) regardless of the fault information (meshing frequency and its harmonics and sideband gear frequencies). HT exalted the area of the low-frequency components, whereas WTK fortified

the high-frequency components in the frequency spectrum (Figure 2.12b,d). WBF was better than the HT and WTK methods because it filtered the noise in some of the meshing frequency harmonics and sideband gear frequencies, but it also reduced or removed significantly informative frequency components (Figure 2.12c). The outANR signal (Figure 2.13), the output signal from the ANR-GRS module proposed here, fulfilled the needs of signal processing: reducing the noise components and preserving the original informative components. It made the vibration signal from the gearbox "cleaner" (lowered the noise) and approached the characteristics of the gearbox vibrations signal presented in Section 2.2.1. This comparison verifies that our accurate is a suitable technique for reducing the noise in gearbox vibration signals and returning an honest reflection of the health states of a gearbox along an electronic signal path.

(a)

(b)

(c)

(d)
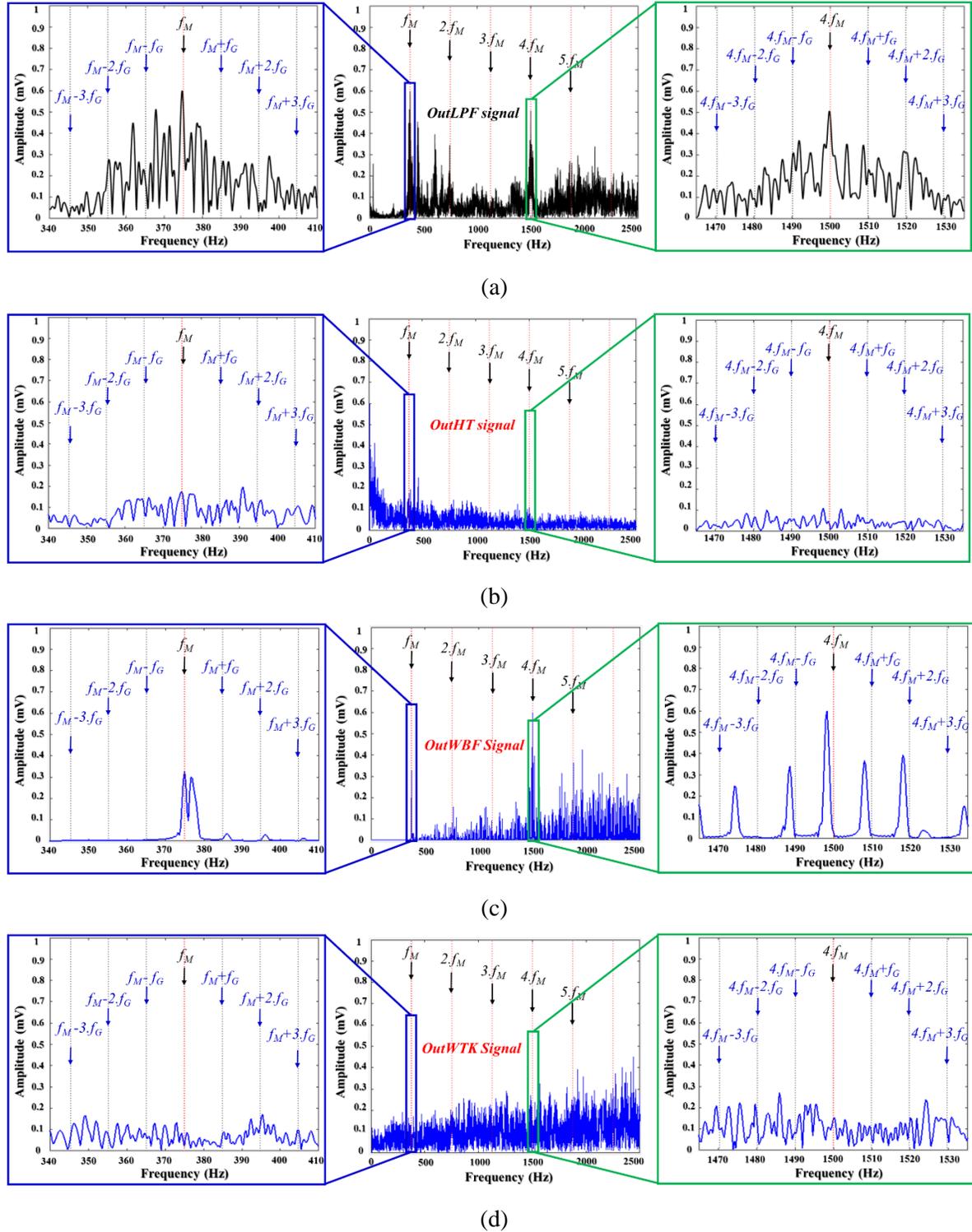
Figure 2.12. The frequency spectrum analysis for the state-of-the-art methodologies:(a) the input signal, (b) the output signal from the Hilbert transform module, (c) the output signal from the window bandpass filter module, and (d) the output signal from the wavelet transform WTK module

47
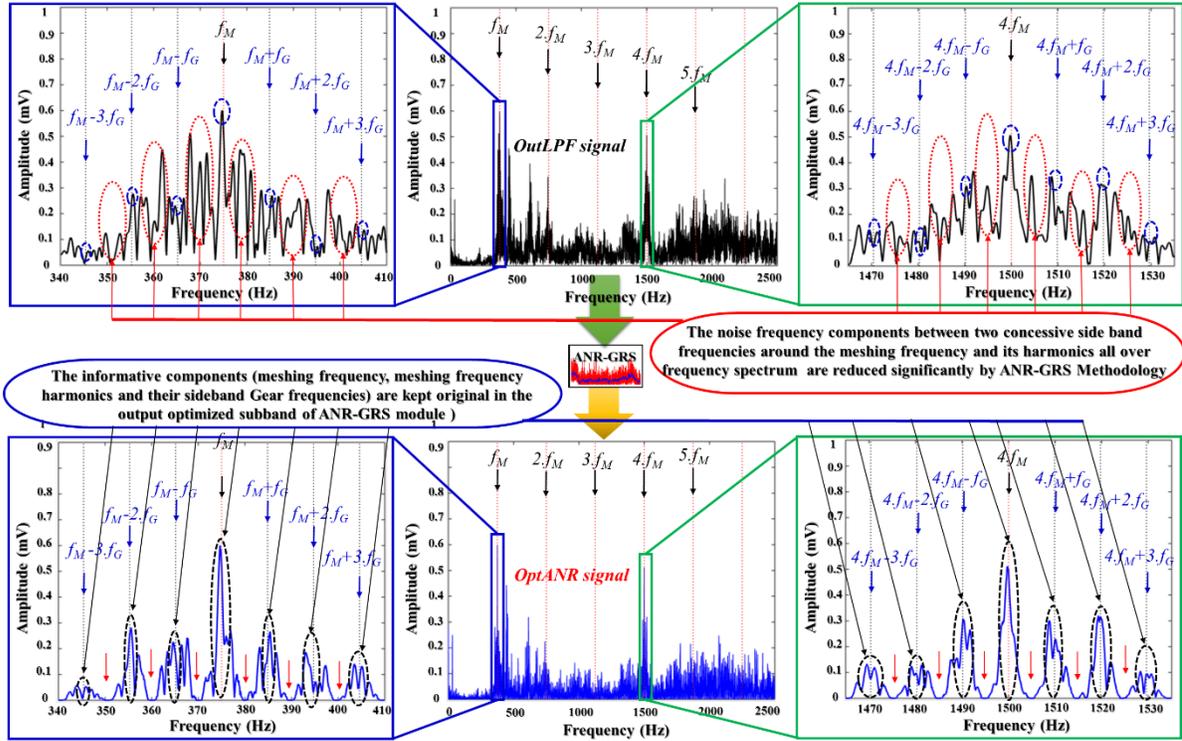
Figure 2.13. Frequency spectrum analysis of the input and output signal of the ANR-GRS module.

2.4.2 Classification results

The classification performance of the proposed methodology is evaluated in two experiments. At the beginning, the feature sets of the dimensions $4 \times 300 \times 21$ (300 samples from each of four health states, as shown in Table 2.2) for each speed (in this experiment, 300 RPM, 600 RPM, 900 RPM, and 1200 RPM) were selected. Then, in the first experiment, the dataset is created by merging the data of four available health states under different rotating speeds resulting in the new feature set of the dimensionality 4800×21. This dataset was then randomly divided into a train and a test sets at ratio 8:2 for training and testing OAOMCSVM classifier and general evaluation of the proposed fault diagnosis methodology.

Furthermore, to prove the robustness of the proposed ANR-GRS technique, the second experiment is performed where the classifier is trained on data (i.e., feature sets) corresponding to single rotating speed and tested by data instances collected under other speeds. Specifically, feature set for a single rotating speed was used as a training set (for instance, feature set corresponding to 300 RPM with the dimensionality of $4 \times 300 \times 21$), and the feature sets corresponding to two other speeds were used for testing (for instance, feature

set corresponding to two other speeds 600 RPM and 900 RPM with the dimensionality of $4 \times 600 \times 21$).

Those processes were run four times. To construct the training model for classification, k-fold cross validation (k-cv) was used to estimate the accuracy of the generalized classification [125]. In k-cv, the set of samples in the feature vector is split randomly into k mutual folds (k=10 in this study), denoted as C1, C2,…, Ck. The classification OAOMCSVM operates on k-times of the accuracy estimation. Some folds {Cj } (a random subset from k folds) are used as a training set, and the rest are used as a validation set and alternative iteration k times. More specifically, for each speed, 300 feature vectors for each health state in the training set were partitioned into ten folds (each fold containing 30 randomly chosen feature vectors (30x21) for each health state); 9 of those folds were used for training, and the 1 remaining fold was used for validation. That process was repeated 10 times until all folds had been used as the validation set. The final measure of performance in the training model is the average value of the accuracies attained in each fold. These data are then used as the testing dataset (which was not used at all in the training process) to verify the OAOMCSVM method and provide the final classification result.

We also used the OAOMCSVM classification method to classify the feature pool configuration datasets extracted from the comparison signal processing methodologies: the raw vibration signal (lowpass filter output signal) extraction (methodology I), HT, (methodology II), WBF (methodology III), WTK (methodology IV), and the IMFs and residuals from the EMD (methodology V). The implementation of those methodologies for achieving the classification result for the four health states complied strictly with the conditions used with input from the ANR-GRS module just described. To estimate the classification result between methodologies (the proposed method and others), all twenty-one features of the vibration signal were used as input feature vectors for the OAOMCSVM module to ensure that the most informative features for and from each methodology were used fairly for the classification. The classification results of the state-of-the-art methodologies and the proposed ANR-GRS methodology obtained during two experiments are shown in Table 2.4 and Table 2.5 and Figure 2.14 to visualize the results tabulated in Table 2.5. Those classification accuracies were computed as follows:

$$C_{accuracy} = \frac{\sum_L N_{TP}}{N_{samples}} \cdot 100\% \qquad (2.23)$$

where L is the number of categories ($L = 4$ as four health states), N_{TP} is the number of true positives (the number of fault samples in category i that are correctly classified as class i), and $N_{samples}$ is the total number of samples used to estimate the performance of the proposed methodology.
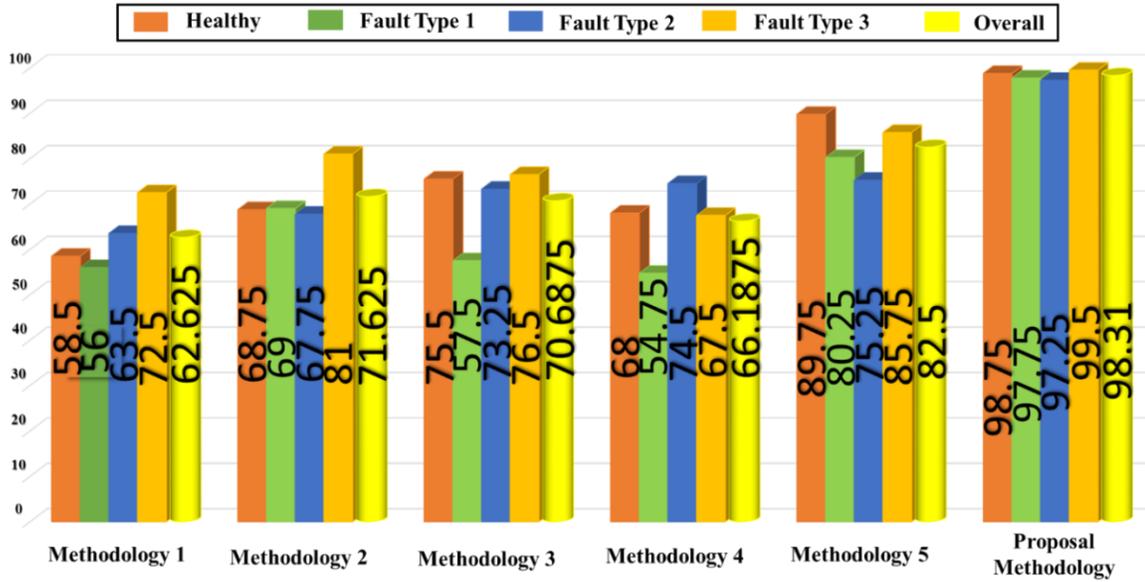
Figure 2.14. The accuracy of each class and the average accuracy of the state-of-the-art methodologies and the proposed ANR-GRS methodology.

Table 2.4 Classification results for state-of-the-art methodologies and the proposed ANR-GRS methodology by a combination dataset of different speeds.

| Methodology | OAOMCSVM (4800 samples) | | Accuracy (%) | | | | |
	Training Set (80%)	Test Set (20%)	Healthy	Fault Type 1	Fault Type 2	Fault Type 3	Overall (%)
I	3840	960	59	73	69	75	69.0
II	3840	960	84	80	67	83	78.30
III	3840	960	92	89	76	83	84.6
IV	3840	960	85	87	58	74	73.10
V	3840	960	92	89	88	94	90.80
ANR-GRS	3840	960	100	99	99	100	99.70

Table 2.5. Classification results for the state-of-the-art methodologies and the proposed ANR-GRS methodology by observation of separated speed dataset.

Metho-dology	OAOMCSVM (10-fold CV)		Accuracy (%)				
	Training set (300 samples)	Test set (600 samples)	Healthy	Fault type 1	Fault type 2	Fault type 3	Overall (%)
I	300 RPM	600RPM, 900 RPM	53	78	69	52	63
	600 RPM	900RPM, 1200 RPM	74	47	53	80	63.5
	900 RPM	600RPM, 1200 RPM	54	46	64	81	61.25
	1200 RPM	300RPM, 600 RPM	53	53	68	77	62.75
	Overall by health states		58.5	56	63.5	72.5	**62.63**
II	300 RPM	600RPM, 900 RPM	51	99	63	85	74.5
	600 RPM	900RPM, 1200 RPM	75	67	64	72	69.5
	900 RPM	600RPM, 1200 RPM	75	48	70	83	69
	1200 RPM	300RPM, 600 RPM	74	62	74	84	73.5
	Overall by health states		68.75	69	67.75	81	**71.63**
III	300 RPM	600RPM, 900 RPM	75	58	69	93	73.75
	600 RPM	900RPM, 1200 RPM	74	70	80	84	77
	900 RPM	600RPM, 1200 RPM	70	49	72	63	63.5
	1200 RPM	300RPM, 600 RPM	83	53	72	66	68.5
	Overall by health states		75.5	57.5	73.25	76.5	**70.69**
IV	300 RPM	600RPM, 900 RPM	64	74	87	63	72
	600 RPM	900RPM, 1200 RPM	82	49	72	64	66.75
	900 RPM	600RPM, 1200 RPM	63	47	69	76	63.75
	1200 RPM	300RPM, 600 RPM	63	49	70	67	62.25
	Overall by health states		68	54.75	74.5	67.5	**66.19**
V	300 RPM	600RPM, 900 RPM	77	94	72	89	83
	600 RPM	900RPM, 1200 RPM	90	82	91	82	86.25
	900 RPM	600RPM, 1200 RPM	94	80	69	85	82
	1200 RPM	300RPM, 600 RPM	98	65	69	83	78.75
	Overall by health states		89.75	80.25	75.25	84.75	**82.5**
ANR-GRS	300 RPM	600RPM, 900 RPM	100	95	98	100	98.25
	600 RPM	900RPM, 1200 RPM	98	99	99	100	99
	900 RPM	600RPM, 1200 RPM	98	99	97	99	98.25
	1200 RPM	300RPM, 600 RPM	99	98	95	99	97.75
	Overall by health states		98.75	97.75	97.25	99.5	**98.31**

Table 2.4 illustrates that the proposed technique significantly outperforms its counterparts when it is trained on the data instances corresponding to all available speeds and achieving the highest accuracy of 99.7%.

Table 2.5 demonstrates that the proposed approach using ANR-GRS also yielded the highest average classification accuracies (98.31%) in comparison with the other five state-of-the-art

signal processing methodologies when it is trained and validated on datasets corresponding to separate rotating speeds.

The methodology I extracted the feature vectors of all four speeds for classification by the OAMCSVM directly from the raw vibration signal (OutLPF signal), in which non-linear and non-stationary signals drown out the informative signal. Accordingly, those results are distributed chaotically among the four classes, producing the lowest accuracy among the 6 methodologies (62.63%). For methodologies II, III, and IV, the vibration signals change with the different characteristics of the gearbox vibration signal (its amplitude and phase modulation signal), so their classification accuracy is also low, around 70%. Methodology V (the EMD technique) is outstanding in comparison with the first four approaches (82.5%) because it extracts IMFs, which contain fault-related information to better discriminate between classes. However, IMFs can be mistakenly extracted from noise components, which damaged the accuracy compared with the ANR-GRS technique by around 15%.

In addition, as a quantitative evaluation, we present the space distribution in a 3-dimensional visualization (Figure 2.15) of samples belonging to four classes based on some features extracted from the outLPF signal and the outANRsignal (signals before and after using the ANR-GRS technique, respectively). The features of the outANR signal show better separation and clustering for different health states of the gearbox fault diagnosis experimental scheme. Samples from the same class are more closely clustered, whereas samples from different classes are discriminated and easy to classify. On the contrary, before using the ANR-GRS, the features of different classes overlap, making it difficult to distinguish the fault classes.
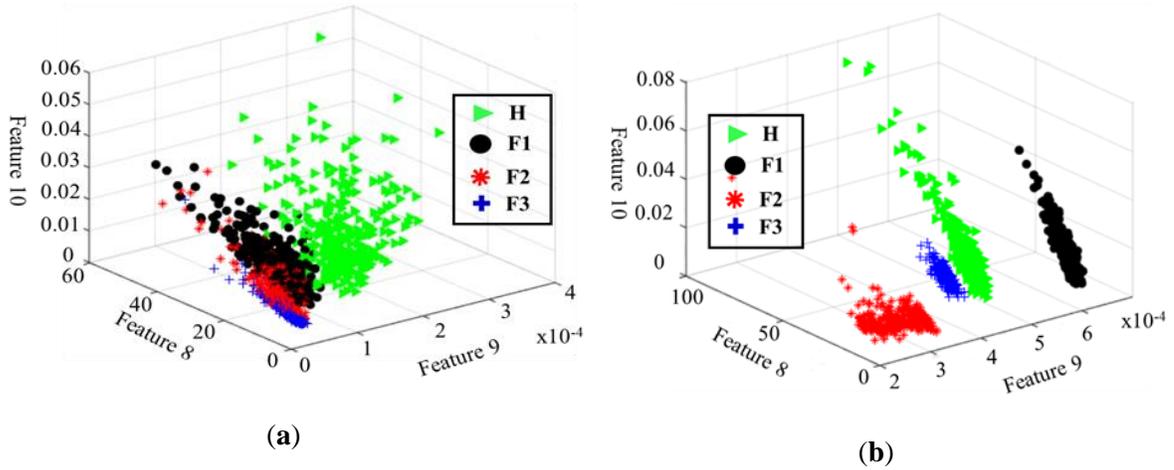
Figure 2.15. Three-dimensional visualization of features extracted from (a) the input signal of the ANR-GRS module and (b) the output signal of the ANR-GRS module.

Moreover, confusion matrixes are shown in Figure 2.16 to demonstrate the reliability of the varying-speed gearbox fault diagnosis methodology using the ANR-GRS module for effective noise reduction. Using real-time tracking of the rotation speed (RPM) of a gearbox system, the ANR-GRS generated speed-related function signals for real-time tracking of speed-dependent noise components, and the optimized output signal was unaffected by speed during classification.

(**a**)

(**b**)

(c)

(d)
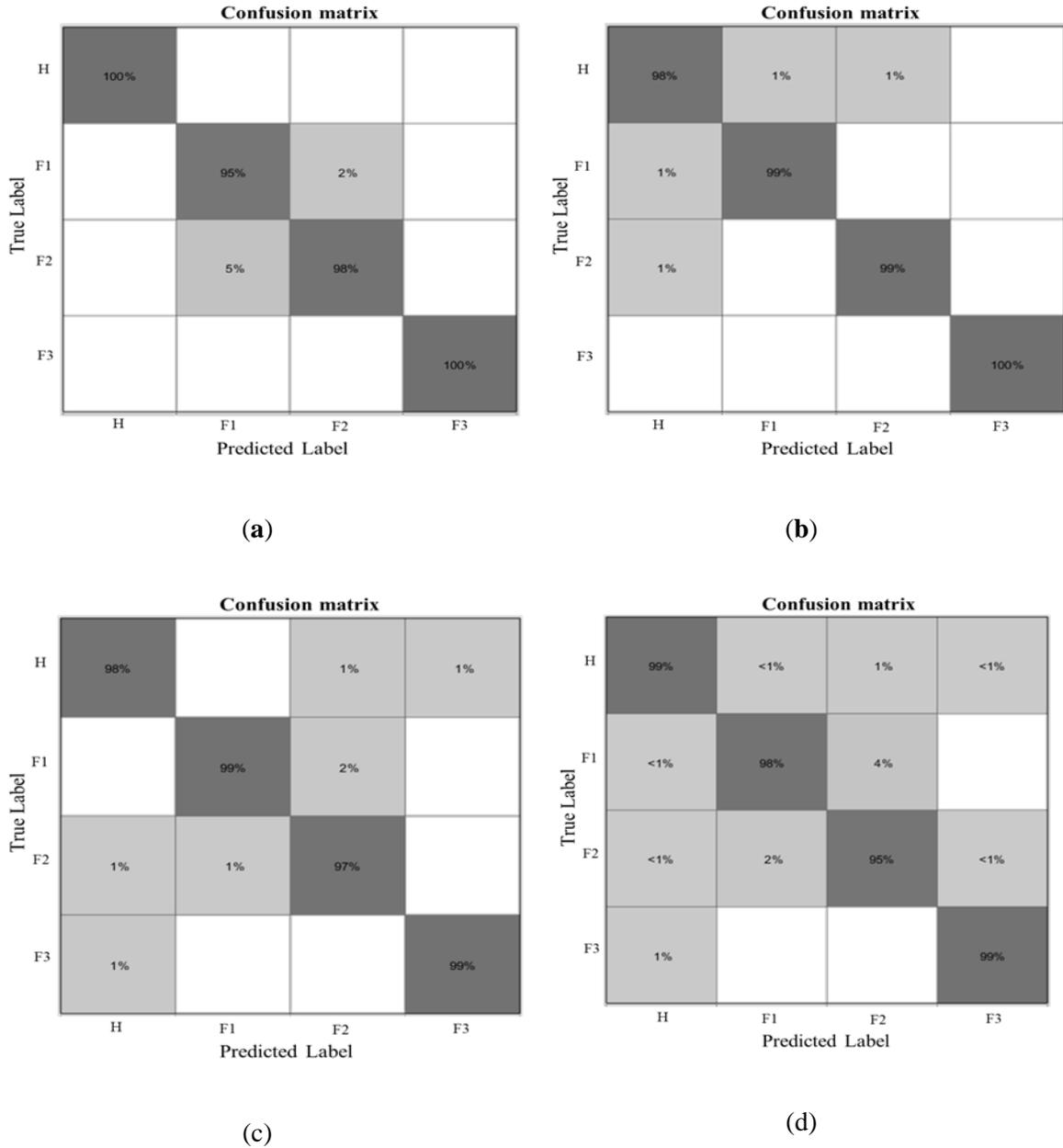
Figure 2.16. Confusion matrixes for four classification cases according to the speed dataset input for training: (a) 300 RPM, (b) 600 RPM, (c) 900 RPM, (d) 1200 RPM.

2.5 Conclusion

In this study we propose a reliable fault diagnosis methodology for gearbox systems under varying speed conditions. It integrates adaptive noise control to significantly reduce noise with

machine learning classification to classify the fault states of the gearboxes. First, we created a set of Gaussian reference signals that are a function of the rotation speed and consist of many noise components such as white noise and band noise that are correlated to the parasitic noise in the vibration signals and independent of the intrinsic informative components. Then, we applied those GRSs to an adaptive noise control technique that produced an optimal sub-band as output for each 1-s vibration sample. The most optimal sub-bands were then used in the feature pool configuration to extract feature vectors, and an OAOMCSVM was used for classification. The experimental results indicated that the proposed gearbox fault diagnosis methodology achieved the highest classification accuracies in both experiments that are equal to 99.7% and 98.31% while significantly outperforming the counterpart state-of-the-art methodologies used for the comparison. In future research, we will continue improving the robustness of the proposed methodology and investigate it's applicability to the real-time fault diagnosis scenarios.

Chapter 3

Fault Identification of Multi-Level Gear Defects using Adaptive Noise Control and a Genetic Algorithm

3.1 Introduction

Gearbox fault diagnosis has been studied excessively and by analyzing its vibration characteristics it is possible to detect the hidden fault-related informative components [82][3]. Vibration and acoustic signals are the two principal techniques for sensing the vibration characteristic of a gearbox. However, the vibration signal is most frequently used due to its easy data acquisition setup [84]. Nevertheless, there are several parasitic noise components observed in the vibration signal that appear from various sources. The most common causes of noise are the resonance processes ongoing in the shaft, gears, and other mechanical components as well as the bias hidden in data collection systems [85]. These factors cause difficulties when attempting to extract fault-related components. To overcome these issues, various techniques addressing the vibration signal processing and feature engineering areas have been proposed by researchers to improve the fault diagnosis capabilities of existing techniques.

Regarding signal processing, many studies focusing on approaches that analyze the signal in multi-domain have been introduced for gearbox fault diagnosis. For instance, those techniques include Hilbert transform with bandpass filters for envelope analysis [91], wavelet transform-based decomposition [99], Hilbert-Huang transforms using time adaptive empirical mode decomposition (EMD) for decomposing vibration signals into intrinsic mode functions (IMF) [104], and a combined technique that utilizes wavelet transform and EMD [110]. Those methods could reduce the signal noise to some degree, but

this noise reduction led to degraded magnitudes of the sideband and meshing frequency harmonics distorting the components of vibration signals related to gear faults. Due to this drawback, these methods are likely inapplicable for differentiating fault types of multilevel gear tooth cut (MGTC) defects. Therefore, in this paper, we apply the adaptive noise control (ANC) for denoising the vibration signals and preserving the original gearbox fault-related components in it [126].

Regarding the feature selection, this procedure from the feature engineering area is mainly used to select the most discriminative feature parameters from the feature set. In general, the feature set extracted from the vibration signals consists of both useful and low-quality features that can affect the classification accuracy of the classifier. To address this problem and remove the redundant features, a searching procedure for selecting the most distinguishable fault features (MDFF) is needed. In practice, several selection algorithms such as: independent component analysis (ICA) [127], principal component analysis (PCA) [128], a linear discriminant analysis (LDA) [129], and a genetic algorithm (GA) are used for selecting high-quality feature. The main performance of these approaches [127]–[129] is to find the important components (independent, principal or linear dependence) in the feature space by applying the statistical process or the statistical functions, it is useful when the fault-related informative components are in the relation with the space of important components, otherwise, the output can be lost fault useful components. GA was constructed based on the principles of natural generic systems and popularly used to output the result of high effectiveness [100], thence GA draws the attention of the researchers. In this paper, a GA implements a heuristic search algorithm on the original feature pool to select the most relevant and discriminative features related to MGTC defects, so the dimensions of feature vectors can be reduced which complements the accuracy of the fault classification process. In other words, GA accommodates a balance between the complexity of computation and the optimal selection. Finally, the MDFF subsets delivered by GA are inputted to the k-nearest neighbor (k-NN) classification algorithm to discriminate the health states of a gearbox system.

The remainder of this paper is organized as follows: Sect. 2 provides a problem

statement of vibration characteristic of a healthy and a defect gearbox. The detail of the proposed methodology is presented in Sect. 3. Section 4 presents the experimental results and discussion, and the concluding remarks are provided in Sect. 5.

3.2 Problem Statements

In this paper, the operating faults of a one-stage transmission gearbox, which encompass a pinion wheel (on drive side) and a gear wheel (on non-drive side) are considered. For the healthy gear the operation of which generates a linear and periodical vibration signal [81], the vibration signal $x_P(t)$ under constant load speed is constructed by using the formula below [121], and the example of frequency spectrum illustrated in Figure 3.1a:

$$x_P(t) = \sum_{i=0}^{M} X_i \cos(2\pi i f_m t + \xi_i), \quad (3.1)$$

where, ξ_i and X_i are the phase and amplitude of the i-th meshing frequency harmonics ($i = 1, ..., N$); f_m is the meshing frequency ($f_m = N_P.f_p$, where f_p is a pinion rotational frequency, N_P is the number of teeth mounted on the pinion wheel; or $f_m = N_G.f_g$ where f_g is a gear rotational frequency, N_G is the number of teeth mounted on the gear wheel); M is the total number of meshing frequency harmonics in the frequency range of a vibration signal.

In the case of faulty gear, the vibration signal, which is directly related to rotating acceleration, is non-linear and non-stationary (i.e. amplitude and phase-modulated signal) [14]. The Eq. (3.2) represents the vibration signal of the defected gear and its spectrum is depicted in Figure 3.1b:

$$x_d(t) = \sum_{l=0}^{M} X_l \big(1 + s_l(t)\big) \cos\big(2\pi l f_m t + \xi_l + \psi_l(t)\big). \quad (3.2)$$

Here: $s_l(t) = \sum_{k=0}^{M} S_{lk} \cos(2\pi k f_g t + \Omega_{lk})$; $\psi_l(t) = \sum_{k=0}^{M} \Phi_{lk} \cos(2\pi k f_g t + \sigma_{lk})$, S_{lk} and Φ_{lk} are the amplitudes and Ω_{lk}, σ_{lk} are the phases of the k-th sideband frequency in the modulated signal around l meshing harmonic.

(a)   (b)
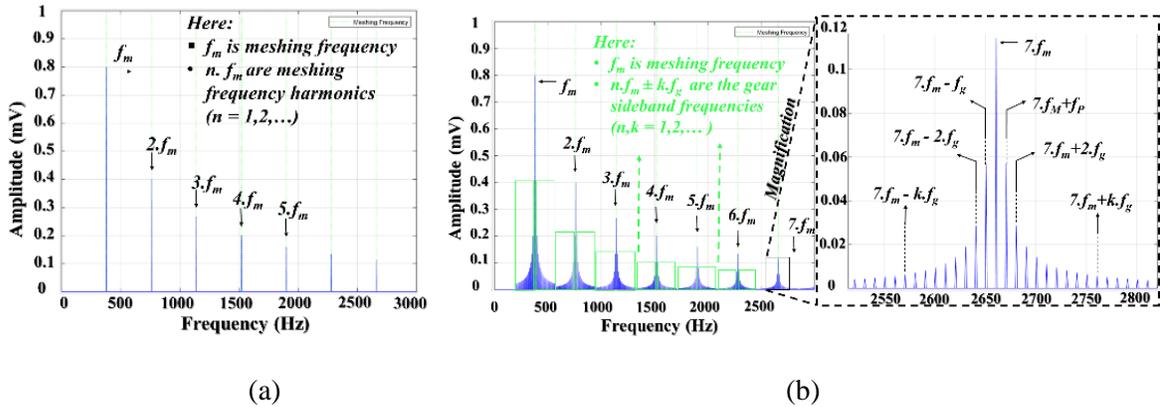
Figure 3.1. The example of the frequency spectra corresponding to the vibration signal of (a) a healthy gearbox and (b) a defected gearbox.

3.3 Proposed Method

The proposed methodology is described in Figure 3.2. It consists of five steps that are represented as follows: the vibration signal collection, adaptive noise control for signal preprocessing, feature extraction, feature selection based on GA, and the k-NN-based fault classification.
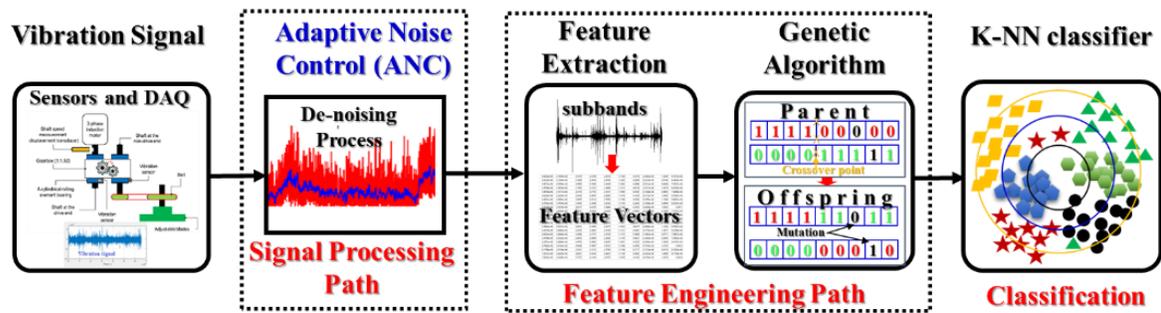
Figure 3.2. Function block diagram of the proposed methodology.

3.3.1 Adaptive Noise Control (ANC)

This paper employed the adaptive noise reducer based Gaussian reference signal (ANRGRS) which was proposed by Nguyen et al. in [126] as an adaptive noise control technique. This technique comprises four main processing modules: a digital filter, an adaptive algorithm, a reference signal generation, and a Gaussian parameter optimizer. The reference signal generation module creates the reference signal inputting to adaptive filter

by combining the adjustable Gaussian (the mean value and standard deviation value are adaptably tuned) and white noise signals which are mostly homologous to the band noise and white noise in a vibration signal, respectively. The adjustable parameters of a Gaussian reference signal (GRS) are the functions of the defective wheel frequency (f_{DW}, which is proportional to a shaft rotational speed). The detail description of GRS is presented by the Gaussian window in the Eq. 3, and the flow chart diagram in the Figure 3.3 [126]:

$$W_{Gau}(s) = \sum_{s=1}^{N_g} e^{-\frac{(s - F_C)^2}{2\sigma^2}}, \tag{3.3}$$

where σ and F_C are the standard deviation and mean value of the Gaussian window (by linearization of the Gaussian function, $\sigma \cong 0.318 \cdot F_C$), respectively. These are the functions of the defective wheel frequency ($F_C = \xi \cdot f_{DW}$ and $\sigma \cong 0.318 \cdot \xi \cdot f_{DW}$). $N_g = \frac{2N_T}{f_s} \cdot f_{DW}$ is the total number of frequency bins in a sideband segment of the gearbox vibration signal with N_T representing the number of samples and sampling frequency f_s. To qualify the homologous condition defined in [126], the parameters of the Gaussian window are adaptively tuned in the following ranges:

$$0.25. f_{DW} \leq F_C \leq 0.75. f_{DW}, \tag{3.4}$$

$$\sigma = \begin{cases} 0.318 * \xi * f_{DW} & when\ 0.25 \leq \xi \leq 0.5 \\ 0.318 * (1 - \xi) * f_{DW} & when\ 0.5 < \xi \leq 0.75 \end{cases}. \tag{3.5}$$

The digital filter module, which is used in the ANC, is an M-order FIR filter with the weight vector as $\mathbf{w}(n) \equiv [w_0, w_1, \ldots, w_{M-1}]^T$. The adaptive algorithm adjusts the weight vectors based on the convergence condition by least mean square (LMS) of error between the vibration signal and each filtered GRS (the reference signal generation module creates a set of GRS signals by adjusting σ and F_C) for fetching the optimal weight vector and local optimal subband signal in the output. The Gaussian parameter optimizer module collects all the local optimal subbands where each of them corresponds to a specific GRS signal to make the set of optimal subbands, then it selects the optimal subband that has a minimum mean square value as global optimal subband termed as an optimized subband, the output of ANC. The block

diagram and signal processing flow are shown in Figure 3.3.
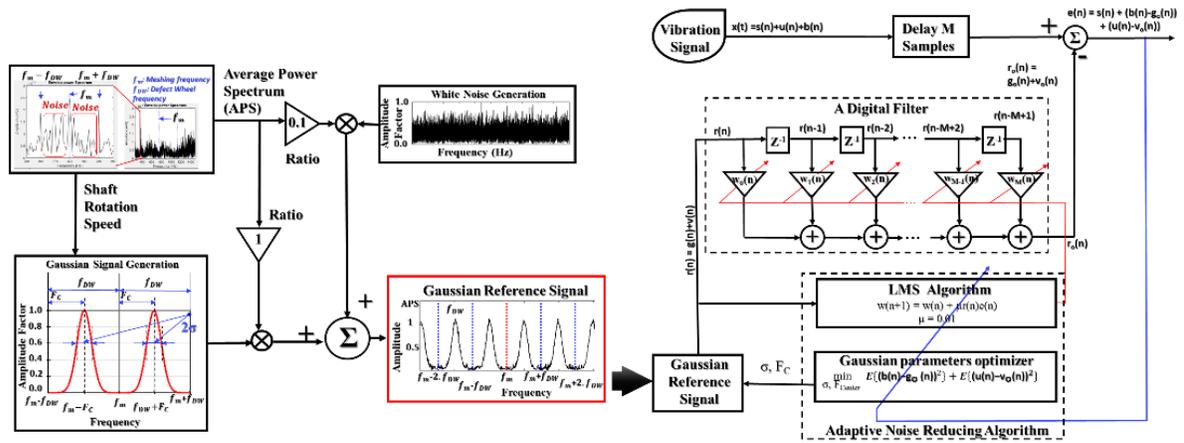
Figure 3.3. A reference signal generation and function block diagram of the ANC module.

3.3.2 Feature Extraction

The optimized subband output from the ANC module represents a reduced-noise vibration signal which carries the intrinsic gear defective information of the MGTC faults. These outputs from ANC are then used to extract eighteen statistical feature parameters from the time and frequency domains to configure the feature pool [124]. The Table 3.1 illustrates the eighteen features with the specific name and calculated formulas [130]. From time-domain, the extracted features are as follows: peak (f1), root mean square (f2), kurtosis (f3), crest factor (f4), impulse factor (f5), shape factor (f6), entropy (f7), skewness (f8), square mean root (f9), energy of system (f10), mean (f11), shape factor of square mean root (f12), margin factor (f13), peak to peak (f14), kurtosis factor (f15). The features extracted from the frequency-domain are as follows: frequency center (f16), power spectral density or frequency spectrum energy (f17), and root variance frequency (f18).

Table 3. 1. Definition of time and frequency domain statistical features.

Features	Equations	Features	Equations	Features	Equations
f1	$Max(|s|)$	f6	$\dfrac{S_{rms}}{\frac{1}{N}\sum_{n=1}^{N}|s_n|}$	f11	$\dfrac{1}{N}\sum_{n=1}^{N} s_n$

f2	$\sqrt{\dfrac{1}{N}\sum_{n=1}^{N} s_n^2}$	f7	$-\sum_{n=1}^{N} p_n \cdot log_2(p_n)$	f12	$\dfrac{S_{srm}}{\frac{1}{N}\sum_{n=1}^{N}\lvert s_n\rvert}$
f3	$\dfrac{1}{N}\sum_{n=1}^{N}\left(\dfrac{s_n-}{\sigma}\right.$	f8	$\dfrac{1}{N}\sum_{n=1}^{N}\left(\dfrac{s_n-\bar{s}}{\sigma}\right)^3$	f13	$\dfrac{max(s)}{S_{smr}}$
f4	$\dfrac{Max(\lvert s\rvert)}{s_{rms}}$	f9	$\left(\dfrac{1}{N}\sum_{n=1}^{N}\sqrt{\lvert s_n\rvert}\right)^2$	f14	max(s)-min(s)
f5	$\dfrac{Max(\lvert s\rvert)}{\frac{1}{N}\sum_{n=1}^{N}\lvert s_n\rvert}$	f10	$\sum_{n=1}^{N} s_n^2$	f15	$\dfrac{Kurtosis}{s_{rms}^4}$
f16	$\dfrac{1}{N_f}\sum_{f}^{N_f} S(f)$	f17	$\sum_{f}^{N_f} S(f)^2$	f18	$\sqrt{\dfrac{1}{N_f}\sum_{f}^{N_f}(S(f)-FC)}$

here is an input signal (i.e., optimized subband), N is the total number of samples, S(f) is the magnitude response of the fast Fourier transform of the input signal s, N_f is total number of frequency bins, $\sigma = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(s_n-\bar{s})^2}$, and $p_n = \frac{s_n^2}{\sum_{n=1}^{N} s_n^2}$

3.3.3 GA-based feature selection

By applying the procedures from evolution theory such as selection, crossover, mutation, and replacement, GA detects the MDFF based on the class-wise information embedded in the complete feature set. The degree of class discrimination (D_{dst}) is defined in Eq. (3.6) for creating separability among observed classes.

$$D_{dst} = \frac{I_{dst}}{W_{dst}}, \tag{3.6}$$

where, I_{dst} is intercross classes discriminating parameter defining the distance between distinct classes, W_{dst} is the distance of the features inside the same class. As the distance between two vectors, the Euclidian distance $D_{Euclidian}(x,y) = \sqrt{\sum_{i=1}^{m}(x_i - y_i)}$ is used in this paper. The parameter D_{dst} tends to increase with the increase of the numerator value (I_{dst}) and the decrease of the denominator. I_{dst} is calculated based on the average distance of specific feature vector from different classes by the formula:

$$I_{dst} = \frac{1}{F.C.M}\sum_{i=1}^{F}\sum_{j=1}^{C}\sum_{k=1}^{m} D_{Euclidian}(i,j,k), \tag{3.7}$$

Equation (3.8) represents the computation of the average distance of a feature in the same class:

$$I_{dst} = \frac{1}{F.M} \sum_{i=1}^{F} \sum_{j=1}^{M} D_{Euclidian}(i,j). \tag{3.8}$$

Here, F is the total number of features ($F = 18$ in this study), C is the total number of categories or classes ($C = 6$), and M is the number of samples of each category.

The GA operates feature optimizing process following the flow chart shown in Figure 3.4 to find out the features with maximum D_{dst} to select as MDFF.
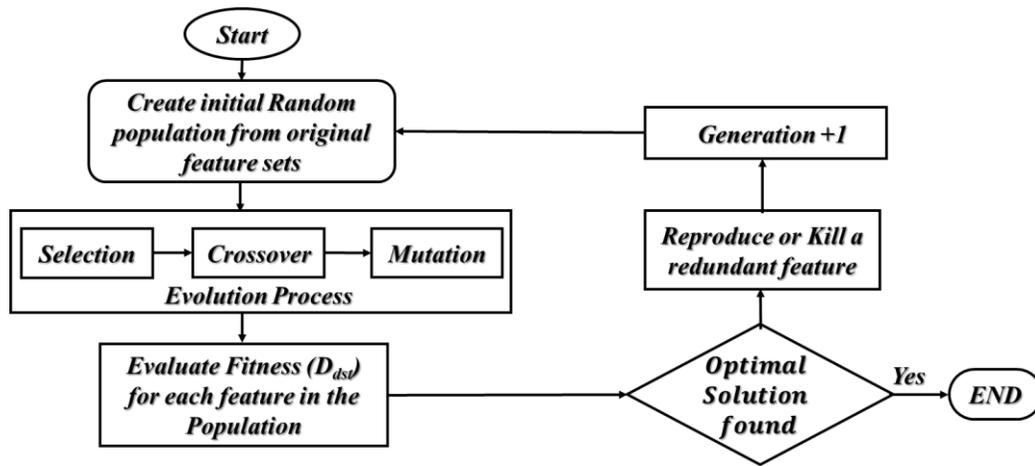
Figure 3.4. The flowchart of the GA-based feature selection

3.3.4 Gearbox Fault Classification using k-Nearest Neighbor Algorithm (k-NN)

In practice, the k-NN algorithm faces problems such as the increase of computational complexity when the high dimensional feature vectors are used as input. Thus, the GA-based feature selection for reducing the dimensionality of the feature set is essential before the application of k-NN to perform fault diagnosis of the gearbox faults. In k-NN, data samples are classed by plurality votes of k-nearest neighbors, which are calculated by distance parameters [131]. Therefore, to complete the k-NN classification process, two important parameters such as the number of nearest neighbors (k) and the distance metric have to be selected. As the distance metric, the Euclidian distance is considered in this study. The appropriate value of k should be selected during the training process. The parameter k could be identified manually or through cross-validation progress. In this paper, the value of k is

first arbitrarily assigned at beginning of the training process, and then its value has been changing during the cross-validation to achieve the optimal value of k.

3.4 Dataset, Experimental Results, and Discussion

3.4.1 Dataset Description

The experimental testbed for the gearbox fault diagnosis system is represented as a spur gearbox that consists of a pinion and gear wheels. The total number of teeth in the pinion wheel is 25 ($N_P = 25$) and 38 teeth ($N_G = 38$) in the gear wheel (the gearbox ratio is 1:1.52), respectively. The tooth failures of the gear wheel have been created by multiple levels of tooth cut seed fault (the total length of the tooth is 9 mm) and were termed as the fault states of the gearbox as follows: tooth cut 10% (D1), tooth cut 20% (D2), tooth cut 30% (D3), tooth cut 40% (D4), tooth cut 50% (D5), and a healthy gear (P). These types of faults are depicted in Figure 3.5. The vibration sensor (accelerometer 622B01) was mounted at the end of the non-drive shaft for sensing the vibration signals of the gear wheel under constant shaft rotation speed at 900 RPM. The analog vibration signals were digitized by using a PCI-2 data acquisition device with a sampling frequency of 65536 Hz. For each fault state of the gearbox, 100 samples of 1-second length were acquired (further referred to as a 1-sec sample).

| No seeded fault | 0.9 mm (10%) | 1.8 mm (20%) | 2.7 mm (30%) | 3.6 mm (40%) | 4.5 mm (50%) |
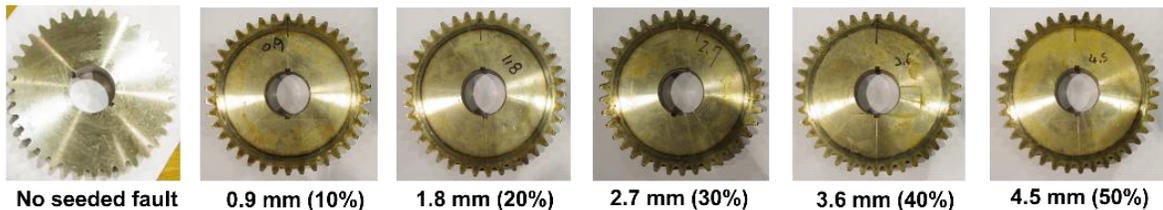
Figure 3.5. The description of the MGTC defects in a gear wheel

3.4.2 Experimental Results and Discussion

As the ANC functions for noise reduction, white noise and band noise integrated into the 1-sec vibration signal samples are used by accessing the spaces between two consecutive sideband frequencies. Then the ANC performs tuning of the model coefficients and the parameters of the Gaussian reference signal to simultaneously remove the noise from the vibration signal and preserve the sideband and meshing frequency harmonics at their original magnitudes. Figure 3.6a presents the effect of applying the ANC technique to one 1-sec

sample of fault state D3. It can be seen that the noise zones in the segment of the frequency spectrum (the green circles) are removed as the multitude in the same segment of the frequency spectrum of optimized subband output from ANC. On contrary, the magnitudes of informative components such as sideband frequencies and the meshing frequency (the black circles) are similar between the input vibration signal and output subband of the ANC.

The GA -based feature selection operates through the selection, crossover, and mutation operators with respect to the finest function of the maximum degree of distance evaluated for each feature in the extracted feature pool. In this experiment, GA is applied during 150 generations to achieve the optimal subset of two MDFF (f9, f17). f9 (square mean root) is a popular feature in the time domain, f17 is a power spectral density or frequency spectrum energy. The distribution zones of 6 MGTC fault states are sketched in Figure 3.6b by a 2D plot with (f9, f17). The subsets of MDFF selected by GA are used as input data of the k-NN classifier to classify the data into their respective categories (i.e., P, D1, D2, D3, D4, D5). By plotting distribution areas of samples in Figure 3.6b, we can see that samples of the same category are close to each other, whereas samples of different categories are separated in feature space.

The optimizing process for tuning k value of the k-NN classifier is implemented multiple times with cross-validation for obtaining the determination boundaries and evaluate the effectiveness of k-NN performance. After optimizing the procedure, the optimal value of k has been assigned as 8. The classification accuracy ($A_{accuracy}$) for assessing the fault diagnosis performance has been calculated as follows:

$$A_{accuracy} = \frac{Number\ of\ True\ Possitive + Number\ of\ True\ Negative}{Total\ number\ of\ Samples} \cdot 100\% \qquad (3.9)$$

By using the optimal k value and GA-based selected MDFF subsets, the achieved an average classification accuracy of the proposed methodology was equal to 97.78%. The confusion matrix obtained by the proposed methodology is depicted in Figure 3.6c. From this figure it can be seen that the classification result for each fault type of MGTC faults was achieved as follows: P (98%), D1 (100%), D2 (100%), D3 (94%), D4 (100%), D5 (94%). From the experimental results, it can be concluded that the pro-posed model is capable of performing

the diagnosis of MGTC gearbox faults at high classification accuracy.

To validate the effectiveness of the proposed method, the comparative experiments are conducted with other conventional feature extraction and classification approaches such as combined feature extraction [126] (further referred to as approach 1), ICA and k-NN (further referred to as approach 2), PCA and k-NN (further referred to as approach 3), and LDA (further referred to as approach 4). These comparisons are implemented by replacing the GA module of the fault identification system for multi-level gear defects in this paper with another one (ICA, PCA, LDA) for selecting the optimal feature space to attain comparing approaches (approach 2, 3, 4). The detail of the comparative results is demonstrated in Table 3.2. It is observed that the approach (GA+k-NN) achieves the highest accuracy (97.78%) outperforming the other state-of-the-art approaches. The improvements of the classification result of the proposed method are 14.26%, 26.37%, 23.51%, and 29.94% in comparison with four referenced approaches: 1, 2, 3, and 4, respectively. It is verified that the combined
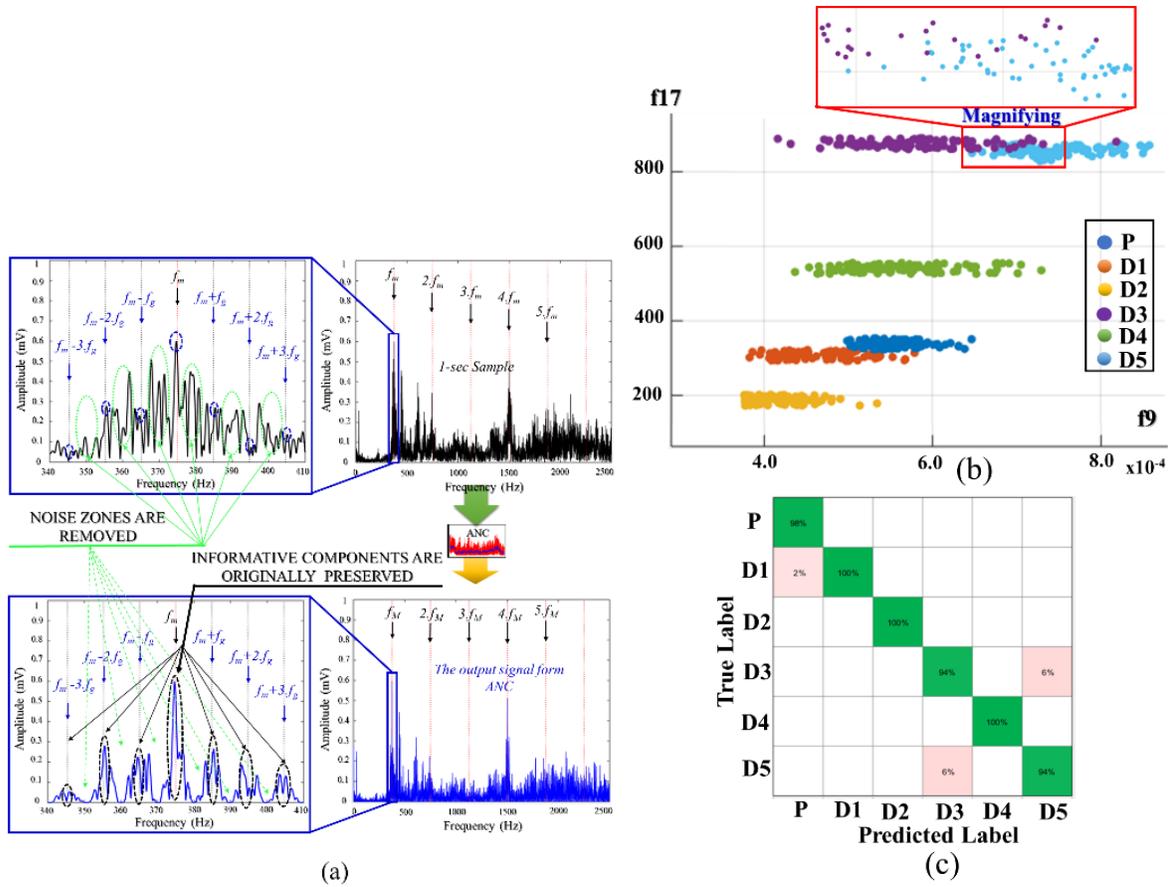
Figure 3.6. The effectiveness of ANC, GA, and classification result: (a) the comparison of vibration signal frequency spectrum between input and output of ANC, (b) 2-D visualization of data samples of MGTC faults based on the MDFF, (c) confusion matrix of the k-NN classification result.

application of ANC and GA can construct an effective model for fault identification of the MGTC gearbox system.

Table 3.2. The fault classification results of the proposed and referenced techniques.

The Approaches	Average Classification Accuracy (%)	The Improvement (%)
Approach 1	83.52	14.26
Approach 2	71.21	26.57
Approach 3	74.27	23.51
Approach 4	67.84	29.94
Proposed Approach	97.78	-

3.5 Conclusions

In this paper, we present the new sensitive gearbox fault identification system for diagnosing MGTC faults with the application of ANC and GA. In combination with ANC, the proposed approach is capable of efficiently removing numerous noise components and simultaneously preserving the intrinsic fault-related components in gearbox vibration signals. The output of GA-based feature selection (MDFF subsets) contains the most discriminative feature parameters that make the samples of each health state being clearly separated in the 2D feature space graph and allow for the application of a simple classification method such as k-NN for discriminating fault categories into the respective classes. The proposed method in this paper yielded the highest average fault classification accuracy result of 97.78% in comparison with conventional approaches. It provides an accuracy improvement of at least 14.26% higher than the referenced techniques. In the future, the proposed method will be investigated for identifying defects of a MGTC gearbox system under varying speed conditions.

Chapter 4

Construction of a Sensitive and Speed Invariant Gearbox Fault Diagnosis Model Using an Incorporated Utilizing Adaptive Noise Control and a Stacked Sparse Autoencoder-based Deep Neural Network

4.1 Introduction

Different types of gearboxes are used in various equipment such as vehicles, industrial machinery, and electrical generators. However, they are prone to defects due to harsh and continuous working conditions. Gear defects can lead to damage of the gearbox sys-tem and become a root cause of damaging the whole mechanical device, which may lead to serious economic losses and the threat of personal safety. Hence, the condition monitoring of gearboxes is essential, and it would be beneficial if the gear defects in gearboxes can be detected in the early stages. The general non-destructive method for condition monitoring of gearboxes is based on sensing the vibration characteristics which contain the fault-related components [80]. The complex sideband frequencies are distributed around the meshing frequency and its harmonics, which are considered as intrinsic components in the vibration signals and are used as informative components to identify gear defects [3], [82]. From the standpoint of signal processing, a gearbox vibration signal is an amplitude and phase-modulated signal that occurs as many frequency tones centered by carrier frequencies are lined up along the whole range of the frequency spectrum. Each set of frequency tones contains a center frequency, as a meshing frequency or its harmonics, and sideband frequencies that are a function of the gear frequencies or specific oscillation frequencies distributed around the

center frequencies. For diagnosing gearbox systems, it is essential to decompose the intrinsic fault-related components, and signal analysis is the most popular technique for these purposes. For capturing the vibration characteristics, accelerometers for measuring vibration signals are more frequently employed than acoustic emission sensors due to their relatively easy installation [83], [84]. Notwithstanding, the vibration signals collected under variable rotational speeds in the gearbox are non-linear and non-stationary signals [132] which accommodate noise caused by the interaction of multiple related systems such as the resonance of shafts, gears, and other mechanical components, electrical and electronic control systems, data collection systems, and the environment [85]. These noise components are random and cause deterioration of the fault-relative characteristics in vibration signals, especially for the vibration signals of MTCG gearboxes (i.e., the noise frequency components might appear randomly with random amplitudes in the whole range of the frequency spectrum of a vibration signal and can cover or deform the original meshing frequency components, its harmonics, and sideband frequencies which are considered as fault signatures). For that reason, the appropriately selected signal processing techniques for reducing noise components and filtering out the informative components are of a high importance.

Recently, many digital signal processing techniques have been developed by re-searchers that can be applied in different domains (e.g., time domain, frequency domain, and time-frequency domain) by employing a variety of advanced approaches such as Fourier transforms, short-time Fourier transforms, Hilbert transforms, wavelet transforms, Hilbert-Huang transform-based empirical mode decomposition [87], [91], [99]–[102], [104], and the combined techniques [105]–[107]. The key methods which were utilized in those methodologies for dis-covering the fault-related components in the vibration signals are as follows: window filtering, thresholding, wavelet excitation, and intrinsic mode function extraction. These methods demonstrated their ability to reduce the noise at some ratio; however, the fault-informative components have been distorted as well. Due to these issues, these methods might not perform well in processing the signals containing MTCG faults to prepare the differentiable data for fault classification. Hence, in this paper, the ANC is utilized for processing the vibration signals to reduce the noise presence and preserve the fault-related components [126] to overcome the disadvantages of the previously introduced signal analysis models.

Considering the feature engineering and classification processes, the traditional gearbox fault diagnosis methods include feature pool configuration (feature extraction and feature selection) and fault classification by machine learning algorithms such as k-nearest neighbors (k-NN), support vector machines (SVMs), and artificial intelligence networks (ANNs) [66], [67], [133], [134]. The main idea of those methods is to perform fault classification using the features which are statistical parameters extracted and selected from vibration signals in the time and frequency domains [58]. Feature extraction is an interfering process that requires a series of experiments for discovering fault-related discriminating feature parameters and then, based on their discriminating capabilities, the appropriated feature selection algorithms are applied for reducing the dimensionality of the constructed feature pool and selecting the most discriminative features for the classification process. These feature pool configuration processes can precede the difficultness of analyzing the vibration signals in each fault case of an MTCG gearbox system for extracting discriminative parameters. Moreover, these approaches can efficiently classify gear faults of a gear-box system under invariant shaft speed, but their performance degrades when applied to the datasets collected under varying shaft speeds. These issues can be addressed by creating a network that can efficiently determine tiny different components of non-stationary vibration signals of an MTCG in a gearbox system operating under varying speeds. The deep learning technique has dawned as an advantageous tool that has been applied in the fields of natural language processing, computer vision, image processing, and pattern recognition, and has succeeded in discriminating barely distinguishable components in categories through multiple non-linear transformations [70], [71], [135]. In other words, deep neural networks (DNNs) are suitable for use in the construction of sensitive and non-linear models. Instead of manually extracting the features and selecting the most separable ones, DNNs can be efficiently used for unsupervised hierarchical feature extraction and feature learning [136]. Thus, this study employs a stacked sparse autoencoder (SSA)-based DNN for identifying the fault types of an MTCG gearbox system based on the vibration signals with reduced noise components delivered by the ANC module.

The major contributions of this study are summarized as follows: (1) an adaptive noise control approach is designed for de-noising and preserving fault-related elements of raw vibration signals to obtain the optimized subbands on its outputs which mostly contain the essential

informative components of vibration signals, and (2) the SSA-DNN utilizes the optimized subbands for identifying the MTCG defect types. The efficiency of the proposed model is evaluated by applying it to the vibration dataset collected from the MTCG gearbox that contains signals collected under six levels of tooth cut fault, such as 6.6%, 10%, 20%, 30%, 40%, and 50% cut as well as signals collected under normal operating conditions. The experimental dataset was collected under variable shaft rotating speeds, such as 300 RPM, 600 RPM, 900 RPM, and 1200 RPM, respectively. The results demonstrate the improved fault classification performance in comparison with the existing models.

The rest of this paper is organized as follows. Section 2 presents a gearbox experimental dataset along with the characteristics of vibration for normal and defective gears. The detail of the proposed method is provided in Section 3. Section 4 describes the experiment configuration and the process of parameter tuning for the proposed network. Section 5 presents the results and discussion, and Section 6 contains the concluding remarks.

4.2 The MTCG Gearbox Dataset

4.2.1 The Experimental Testbed and MTCG Gearbox Dataset

Figure 4.1 shows the experimental setup used for exploring the vibration characteristics of the MTCG gearbox system. A three-phase AC induction motor is connected to a pinion wheel through a drive shaft (DS) and a set of adjustable blades is mounted on a non-drive shaft (NDS) the other end of which is connected to a gear wheel. The numbers of teeth on the pinion
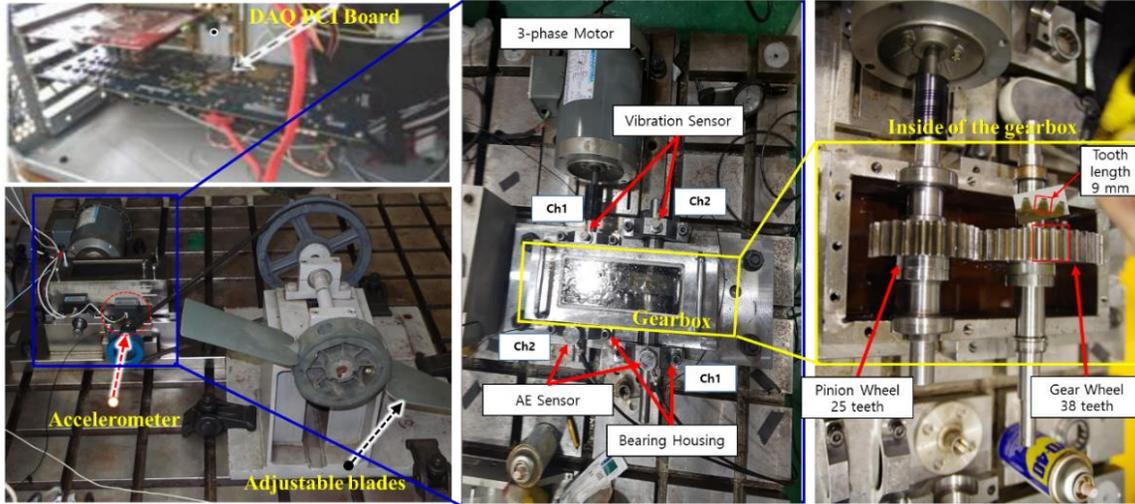
Figure 4.1. Experimental testbed arrangement for acquiring the MTCG gearbox dataset.

(a)   (b)   (c)   (d)   (e)   (f)   (g)
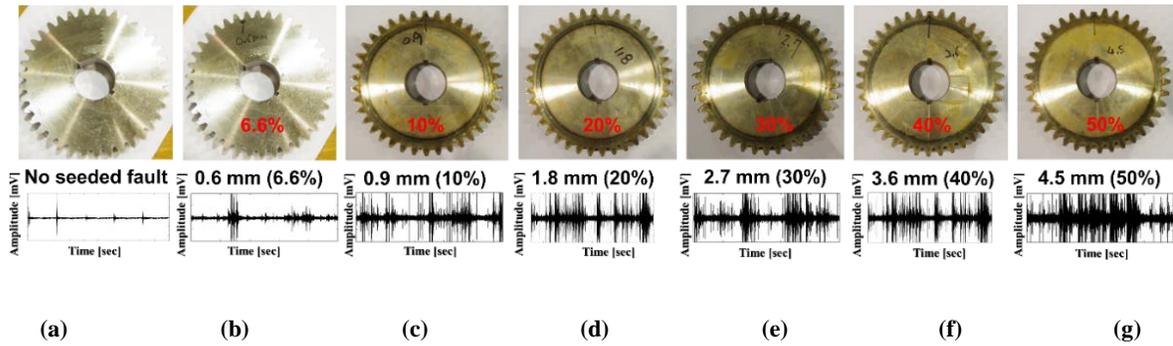
Figure 4.2. The defect states of the gear wheel and examples of vibration signals at a rotation speed of 600 RPM: (a) no seeded fault, normal gear, (b) tooth cut 6.6% (0.6 mm), (c) tooth cut 10% (0.9 mm), (d) tooth cut 20% (1.8 mm), (e) tooth cut 30% (2.7 mm), (f) tooth cut 40% (3.6 mm), and (g) tooth cut 50% (4.5 mm), respectively.

wheel and the gear wheel are equal to 25 ($N_p = 25$) and 38 ($N_g = 38$), respectively. The length of each tooth is equal to 9mm. The torque generated by the AC motor is transferred to the adjustable blade through the gearbox with a gear ratio of 25:38 (1:1.52). The multi-level tooth cut faults were seeded in one tooth of the gear wheel by cutting the percentage of the tooth length as depicted in Figure 4.2. The MTCG fault types contain a normal gear or a no seeded fault gear (N) condition, a tooth cut seeded gear defect of 6.6% (D1), a tooth cut seeded gear defect of 10% (D2), a tooth cut seeded gear defect of 20% (D3), a tooth cut seeded gear defect of 30% (D4), a tooth cut seeded gear defect of 40% (D5), and a tooth cut seeded gear defect of 50% (D6), respectively. For measuring the vibration characteristics of an MTCG gearbox in the normal and defects cases, the vibration sensor (an accelerometer 622B01 of IMI Sensor

company) was installed at the end of the NDS, 72.5 mm from a gear wheel. Therewith, the shaft rotation speeds are monitored by using a displacement transducer (a speed sensor) to track the seeded hole in the DS once per rotation. The output signal from a vibration sensor was digitized using a PCI-based data acquisition board with a sampling frequency of 65,536 Hz continuously for one second. The data collection process was repeated 200 times to receive 200 samples of 1-sec length per each gear defect state (seven states) under each shaft rotation speed. Therefore, the total number of observing samples is 5600, each of one second duration. The detailed description of the MTCG gearbox dataset is in Table 4.1.

Table 4.1. A detailed description of the MTCG defect types and dataset.

Gearbox Fault State.	Description	Number of 1-Sec Data Samples Acquired for Each Rotation Speed				Sampling Frequency (Hz)
		300 RPM	**600 RPM**	**900 RPM**	**1200 RPM**	
Normal Gear (N)	No seeded fault in the teeth of a gearbox	200	200	200	200	65,536
Defect type 1(D1)	Gear tooth cut 6.6% (0.6 mm)	200	200	200	200	65,536
Defect type 2(D2)	Gear tooth cut 10% (0.9 mm)	200	200	200	200	65,536
Defect type 3(D3)	Gear tooth cut 20% (1.8 mm)	200	200	200	200	65,536
Defect type 4(D4)	Gear tooth cut 30% (2.7 mm)	200	200	200	200	65,536
Defect type 5(D5)	Gear tooth cut 40% (3.6 mm)	200	200	200	200	65,536
Defect type 6(D6)	Gear tooth cut 50% (4.5 mm)	200	200	200	200	65,536

4.2.2 The Vibration Characteristics of the Gearbox System

The categories of gear defects can be generally split into three types: manufacturing defects (wheel eccentricity, defect of tooth profile, etc.), installation defects (parallelism), and defects caused by long-term operation (cracked tooth, spalled tooth, case ware tooth, tooth wear, etc.). In this work, the MTCG defects were created to simulate the operated defects as the multi-level depth of a tooth cut seeded in the gear wheel of the gearbox system. The vibration characteristics of a gearbox system are analyzed in the cases of a healthy gear (a defect-free gear) and a defect gear for identifying the informative fault-related components in the vibration signal. The vibration signal of a defect-free gear represents a linear and periodical signal that is calculated using the following formula [121]:

$$y_n(t) = \sum_{k=i}^{K} Y_k \cos(2\pi k f_h t + \partial_k) \tag{4.1}$$

where $y_n(t)$ is a vibration signal of a healthy gear; K is a total number of meshing frequency harmonics in the observed frequency spectrum of a vibration signal; Y_k and ∂_k are the amplitude and phase of the k-th meshing frequency harmonics ($k = 1,\ldots, K$); and f_h stands for the meshing frequency which can be calculated using the parameters of a gear wheel ($f_h = f_g N_g$, where f_g is a gear wheel rotation speed and N_g is the number of gear teeth) or parameters of a pinion wheel ($f_h = f_p N_p$, where f_p is a pinion wheel rotation speed and N_p is the number of pinion teeth). Figure 4.3a illustrates an example of a frequency spectrum denoting the informative components as meshing frequency tones in a spectrum of vibration signals of a defect-free gearbox.

Compared to a vibration signal of a normal gear, a signal of a defected gear is more complex due to the occurrence of impulsive vibrations when the motion is transferred from the DS to the NDS by rotating a pinion wheel through a gear wheel at a defective tooth position during one rotation cycle. Those periodical impulsive vibrations create the non-linear and non-stationary vibration signal formed as the amplitude and phase modulation signal in the point
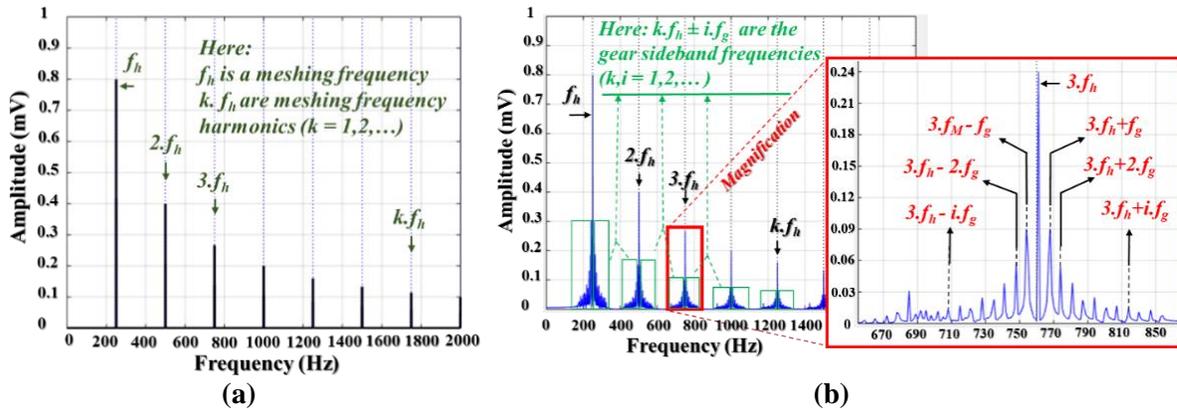
Figure 4.3. The example vibration signals present in the frequency domain: (a) a normal gearbox and (b) a defective gearbox

of view in the signal processing zone [82]. The fault gear vibration signal can be formulated [81] by equation (4.2), and an example for demonstrating the fault-related informative components is shown in Figure 4.3b.

$$y_d(t) = \sum_{k=0}^{K} S_k\big(1 + \sigma_k(t)\big) \cos\big(2\pi k f_h t + \epsilon_k + \psi_k(t)\big) \tag{4.2}$$

Here, $\sigma_k(t) = \sum_{i=0}^{N} \Theta_{ki} \cos(2\pi i f_g t + \Omega_{kj})$ and $\psi_k(t) = \sum_{i=0}^{N} \Psi_{ki} \cos(2\pi i f_g t + \xi_{ki})$ are modulating components of the amplitude and phase partial in the fault gear vibration signal $y_d(t)$; Θ_{ki}, Ψ_{ki} are amplitudes and Ω_{kj}, ξ_{ki} are phases of the i-th sideband, respectively, roundly k-order meshing the frequency tone of the vibration signal $y_d(t)$.

4.3 The Incorporated Construction Model of the ANC and the SSA-DNN

The proposed sensitive and speed invariant model for diagnosing gearbox faults is presented in Figure 4.4. Three major function blocks are utilized in this model, such as the data collection system (Sensors and DAQ), the ANC, and the SSA-DNN. The data collection system collects the vibration dataset of an MTCG gearbox system for each fault type (seven fault types in total) under variable shaft rotation speeds. It collects the vibration data samples and captures the gear defect behaviors in the vibration characteristics: each vibration sample is evenly acquired during one second to monitor several complete rotation cycles of the defected gear. The ANC module then processes the raw vibration signals. Firstly, it performs down-sampling three times along with filtering the signal with a low-pass filter to receive the vibration subbands within the frequency range from 0 to 10 kHz according to the real operating frequency range of the vibration sensor [126]. The expression of multi-level gear defect types on the vibration characteristic is signified by the magnitudes of the principal frequency tones, therefore the main function of the ANC is optimizing vibration subbands for removing the redundant components along with noise while preserving the original fault-related components. The output of the ANC provides the optimized subband in the frequency domain (power spectrum density) which mostly contains the meshing frequency, its harmonics, and their distributed sideband gear frequency tones (i.e., the defect-related informative components). Under variant speeds condition, the positions of principal frequency tones are altered according to the explanation in Section 2. There exist the components that represent the speed invariant MTCG defects as the numbers of latent features related to the ratio and proportional to the amplitudes and displacements in the optimized vibration subbands, which

are difficult to extract features from by traditional methodologies [124]. Notwithstanding, based on the unsupervised learning and hierarchy of feature extraction constitution of a deep neural architecture (DNA), the SSA-DNN can vanquish the issue and automatically explore the most defect-substantial features from a set of components in the frequency spectrums of optimized subbands output from the ANC. By fetching out these features, the SSA-DNN can use them to identify defect types of an MTCG gearbox system for achieving a high classification result in the output layer.

4.3.1 Adaptive Noise Control (ANC)

ANC is a signal processing method used for reducing noise and preserving the fault-related informative elements in gearbox vibration characteristics. The ANC approach is a self-
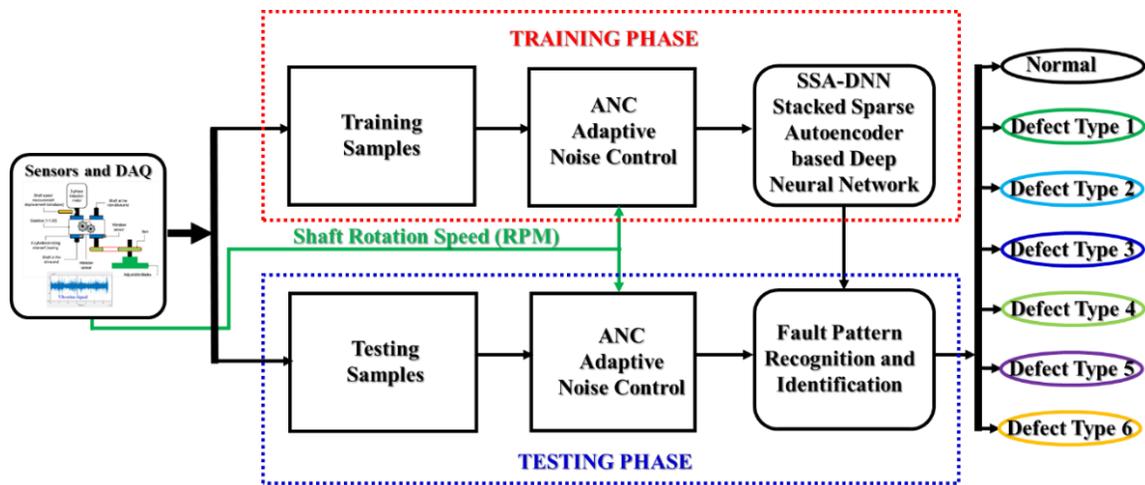
Figure 4.4. A block diagram of the proposed gearbox fault diagnosis model.

constructed and time-varying system that uses a recursive algorithm for optimizing its parameters for obtaining the desired optimized signal in its output [123]. General ANC consists of a digital filter, an adaptive algorithm, and a reference signal generator. An adaptive algorithm operates to update the coefficients of the digital filter based on the feedback error signal of a filtered reference and an input signal to receive the optimized denoised subband signal in the output of the ANC [114]. In this study, the ANC employs the adaptive noise reducer-based Gaussian reference signal (ANR-GRS) which has been elaborated in [126] for reducing noise and optimizing gearbox vibration signals. An adaptive noise control scheme

contains two inputs (the desired input and a reference input) and one output. As the desired input for the observed signal, the vibration subband is used in this study, while the reference input is used for a signal that imitates the parasitic noise in the observed signal. The function of the ANC approach can be described in detail in the following processes [126]:

1) Generating the reference signal to supply to the reference input of an ANC:

Mainly, there are two types of noise present in the vibration signal: white noise and band noise. Hence, the reference signal generator creates the output signal behavior which is homologous with those such as Gaussian signals and white noise signals, as illustrated in Figure 4.5. The parameters of a Gaussian signal (a mean and a standard deviation value) can be adjusted based on the input variable of the shaft rotation speed. The adjustable Gaussian window, a component for building the entire Gaussian signal, is drawn to adapt to the frequency space between two consecutive sideband gear frequencies, formulated as follows:

$$W_{Gref}(p) = \sum_{p=1}^{N_t} e^{-\frac{(p-F_o)^2}{2\sigma^2}} \tag{4.3}$$

where the adjustable parameters (mean value F_o and standard deviation value σ) are functions of the shaft rotation frequency [18]. Concretely, F_o is proportional to the frequency of faulty wheel (f_{DG}) and can be computed as below:

$$F_o = \varepsilon \cdot f_{DG}, \tag{4.4}$$

and by linearizing the Gaussian function, the standard deviation is approximated to the mean value as:

$$\boldsymbol{\sigma} = 0.318 \cdot F_o = 0.318 \cdot \varepsilon \cdot f_{DG}. \tag{4.5}$$

Also, the number of sideband segments $\mathbf{N_t}$ is calculated using the known parameters such as the number of samples $\mathbf{N_s}$, sampling frequency $\mathbf{F_s}$, and fault wheel frequency. The formulation of sideband segments is presented below:

$$\mathbf{N_t} = \frac{2\mathbf{N_s}}{\mathbf{F_s}} \cdot f_{DG} \tag{4.6}$$

where the frequency of a faulty wheel ($\boldsymbol{f_{DG}}$) is represented as a gear frequency ($\boldsymbol{f_g}$) which is defined in Section 4.2. Therefore, by adjusting the ratio coefficient $\boldsymbol{\varepsilon}$, the Gaussian window

can access the space between two consecutive sideband frequencies in the frequency spectrum of a vibration signal to reduce the presence of noise. According to specific conditions defined in [126], first, the coefficient $\boldsymbol{\varepsilon}$ is selected from the range of [0.25 0.75], and then, the Gaussian windows are created with the parameters chosen as shown below:

(1) the mean value F_o is assigned to be in the range:

$$0.25 \cdot f_{DG} \leq F_o \leq 0.75 \cdot f_{DG} \tag{4.7}$$

(2) the standard deviation of the Gaussian windows is selected in the following range:

$$\sigma = \begin{cases} 0.318 \cdot \varepsilon \cdot f_{DG} & when\ 0.25 \leq \varepsilon \leq 0.5 \\ 0.318 \cdot (1 - \varepsilon) \cdot f_{DG} & when\ 0.5 < \varepsilon \leq 0.75 \end{cases} \tag{4.8}$$

By limiting the adjusting values of the coefficient $\boldsymbol{\varepsilon}$, each generated Gaussian window is positioned completely inside the area between two consecutive sideband frequencies during the optimization processes in the next steps. This ensures that the adaptive noise control technique performs reducing band-noise significantly whereas originally preserving the fault-
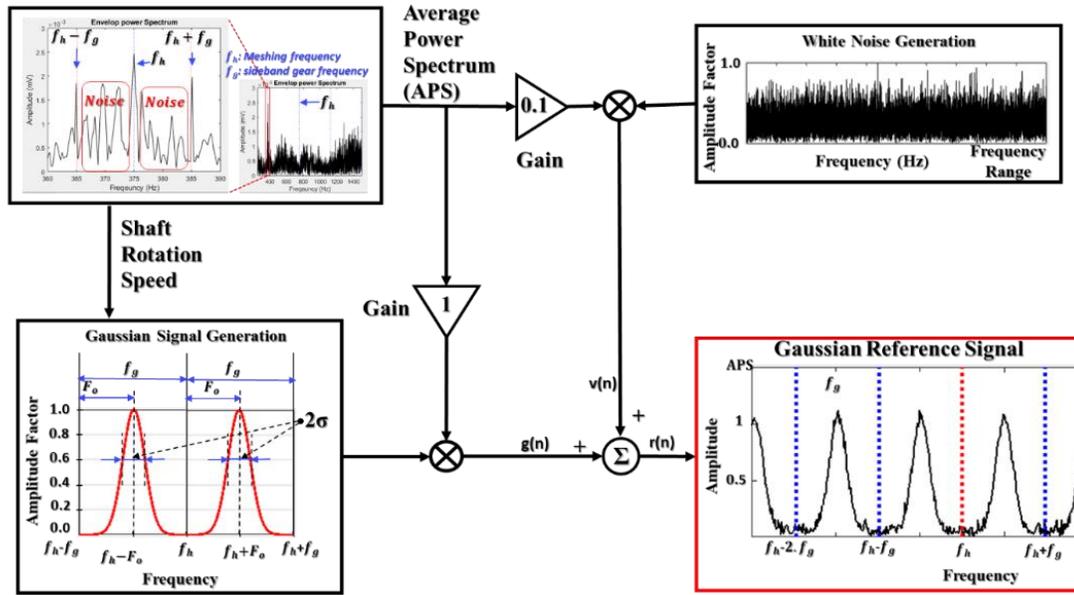
Figure 4.5. A functional block scheme of generating the adjustable Gaussian reference signal.

related informative components as meshing frequencies, its harmonics, and sideband frequencies [126].

2) The construction of an adaptive filter

The adaptive filter is formed by combining the N-tap FIR digital filter (the coefficient vector

as $\mathbf{c}(n) \equiv [c_0, c_1, \ldots, c_{N-1}]^T$) and a least mean square (LMS) adaptive algorithm. The reference signals are used as the input to the digital filter and its output signals are summed with the vibration subbands to calculate the output error signals. Based on this error, the LMS adaptive algorithm tunes the coefficient vectors according to the convergence criterion of the least mean square error for determining the optimal coefficient vector ($\mathbf{c}_o$) and then identifying the local optimal subbands. The operation of an adaptive filter is functionally described in Figure 4.6.
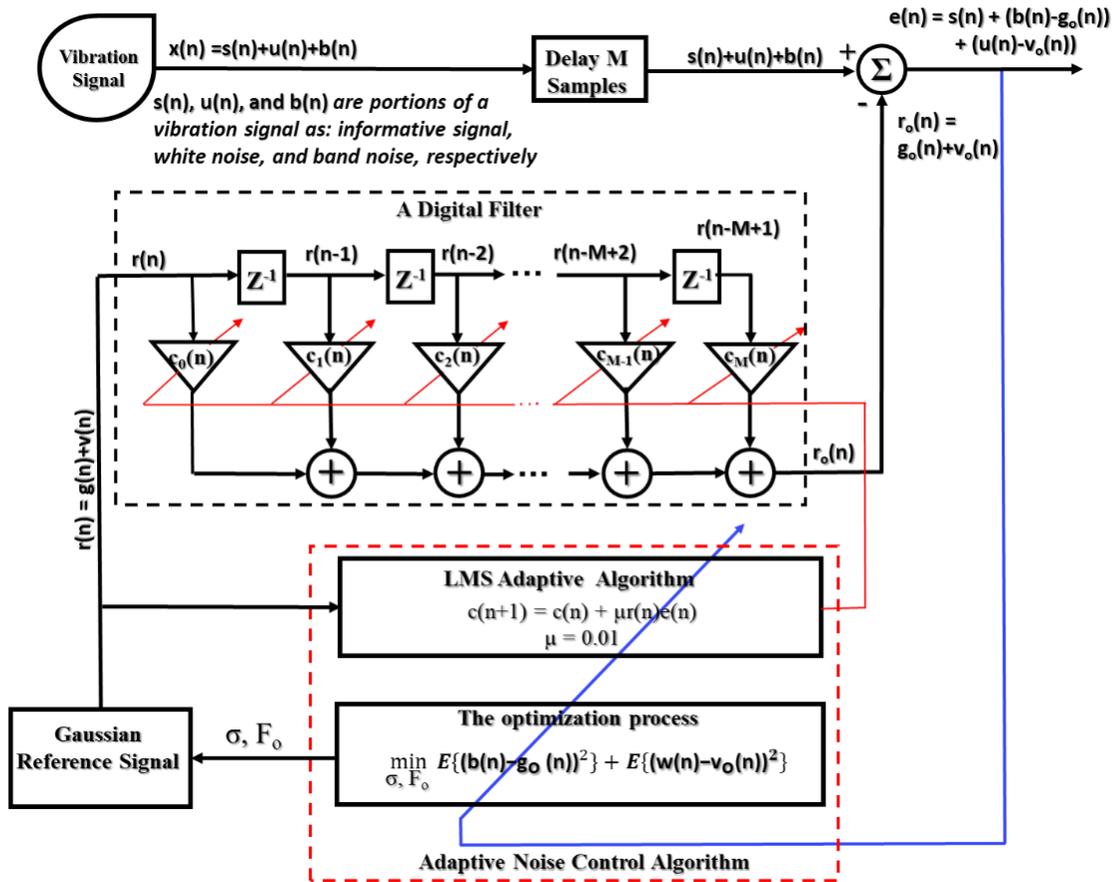
Figure 4.6. A functional block diagram of an adaptive noise control module.

3) The optimization process for selecting the optimal vibration subband

Each vibration subband, processed by an adaptive filter with the input reference of a parameter- adjustable Gaussian reference signal, results in many subbands in its output (termed as local optimal subbands) corresponding to the set of specific values of parameters and appropriate optimal coefficient vectors. At this step, the ANC selects the subband which

has a minimum mean squared value as an output result of the optimization process (termed as an optimized subband) illustrated in Figure 4.6. This optimized output subband is a final output of the ANC module that contains mostly the fault-related informative components and trivial disturbances or redundant components.

In fact, the signal portions, which reflect the gear states (a meshing frequency, meshing frequency harmonics, and gear sideband frequencies), are represented mostly in the frequency domain as magnitudes, tones amplitudes, oscillations, frequencies, and the ratios between them. Thus, it is suitable to use the frequency spectrum of the optimized subband as the input data to the SSA-DNN so the deep network can explore and automatically extract the defect characteristic features from its inputs. Additionally, the usage of the frequency spectrum of the vibration signal reduces the complexity of the DNN. Therefore, in this paper, the frequency spectrum of the optimized vibration subband calculated by Fourier transform [137] is used as the input of the SSA-DNN module. The spectrum of the optimized subband is of ranges from 0 to 10 kHz due to the down-sampling process of raw one second vibration samples.

4.3.2 Stacked Autoencoder

A stacked autoencoder is a type of DNN, with a number of hidden layers greater than one, formed by stacking simple autoencoders for feature discrimination and classification. To understand the concept of a stacked autoencoder, a simple autoencoder should be discussed first. It is an unsupervised DNN based on a three-layer symmetrical architecture for learning the representation of high-level data [75]. An autoencoder functions through two learning stages-encoding and decoding, as shown in Figure 4.7. In the encoding stage, it transforms the higher-dimensional input into a lower-dimensional one. High-dimensional input data is compressed by the hidden layer in DNN architecture [138]. Hence, the encoding path contributes to the principal goal of an autoencoder. In the mathematical expression, the higher-dimensional input represented as $s \in R^N$ (i.e., N dimensions) is encoded to a lower-dimensional space $h \in R^K$ (i.e., K dimensions), producing the output vector known as a latent space. The encoder function or the latent space can be represented as follows:

$$h = f_e(W_e s + b_e), \tag{4.9}$$

where f_e, W_e, and b_e are the encoding activation function, weights, and bias of the network, respectively. From Figure 4.7, it can be interpreted that the decoding portion reconstructs the output of a lower-dimensional space that was compressed from higher-dimensional input using an encoding process. The reconstruction procedure can be expressed as follows:

$$\hat{s} = f_d(W_d h + b_d) \tag{4.10}$$

Here f_d, W_d, and b_d are the decoding activation function, weights, and bias of the network, respectively. The key goal of the autoencoder is to minimize the reconstruction loss which is an objective function of an autoencoder. It can be expressed as following [76]:

$$\mathcal{L}(s, \hat{s}) = \pm(\|s - \hat{s}\|) = \|s - f_d(W_d(f_e(W_e s + b_e)) + b_d)\| \tag{4.11}$$

In this paper, the feature engineering and classification path of the sensitive and speed invariant gearbox fault diagnosis model is constructed by stacking multiple sparse autoencoders as a stacked sparse autoencoder (SSA) for determining the small differences of features between gear defect types which are the basis components for improving classification accuracy. In the next subsection, the sparse autoencoder algorithm is explained.
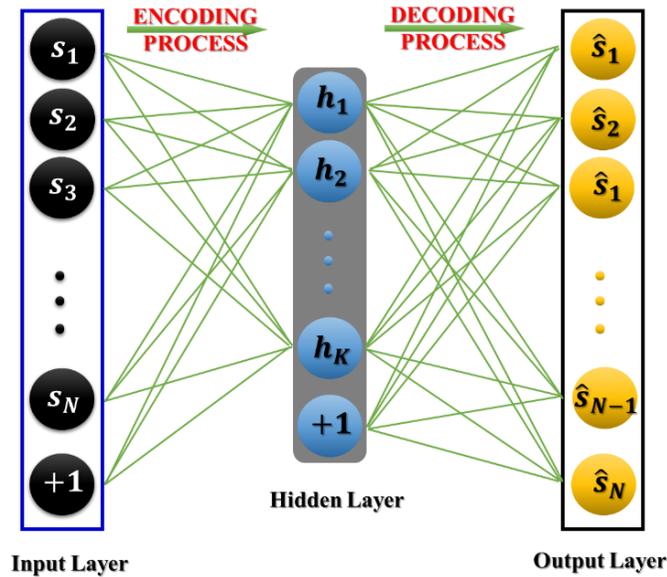
Figure 4.7. The diagram of the two learning processes of an autoencoder.

4.3.3 Sparse Autoencoder

Sparsity is a special parameter of autoencoders, which puts a constraint onto the hidden layer and causes activation of inactive hidden units to discover the tiny differences in decimated

features of data representation more sensitively and robustly than the simple autoencoder architecture [79]. The constraint of a sparse autoencoder usually embeds a regularization term to the objective function. Therefore, the regularized objective function can be expressed as follows [76]:

$$\mathcal{L}(s, \hat{s}) = \frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{K} (s_{kn} - \hat{s}) + \beta \times \phi_{weights} + \gamma \times \Phi_{sparse} \tag{4.12}$$

In Equation (4.12), β and γ refer to the $\mathcal{L}_2$ regularization coefficient and the sparsity penalty factor, respectively. In the training process of an autoencoder, it is sometimes observed that the value of γ alters in an inversed way with the values of weight parameters and behaves proportionally to the latent space h (for example the value of the sparsity penalty factor increases by decreasing the value of weights and increasing the value of latent code). Thus, the $\mathcal{L}_2$ regularization is introduced for embedding in the cost function to solve this issue, which can be represented as follows [76]:

$$\phi_{weights} = \frac{1}{2} \sum_{l}^{L} \sum_{i}^{n} \sum_{j}^{k} (W_{ij}^{l})^2 \tag{4.13}$$

where L, n, and k represent the number of hidden layers, the number of observations, and the number of variables in the input data, respectively. Consequently, the sparsity constraint

$$\Phi_{sparse} = \sum_{i=1}^{L^{(1)}} KL(\rho||\bar{\rho}) = \sum_{i=1}^{L^{(1)}} (\rho \log \frac{\rho}{\rho_i} + (1 - \rho) \log(\frac{1 - \rho}{1 - \rho_i})) \tag{4.14}$$

Φ_{sparse} can be formulated as follows:

where

$$\rho_i = \frac{1}{m} \sum_{j=1}^{m} z_i^1(s_j) = \frac{1}{m} \sum_{j=1}^{m} h(w_i^{(1)T} s_j + b_i^{(1)}) \tag{4.15}$$

This Equation (4.14) is known as Kullback-Leibler divergence [139]. Φ_{sparse} takes a higher value when the i-th neuron gives an average activation value $\bar{\rho}$ because that deviates mainly from the desired value ρ.

To establish the SSA, several numbers of sparse autoencoders, which have been individually trained, are stacked and positioned in a form such the input layer is placed before the series of
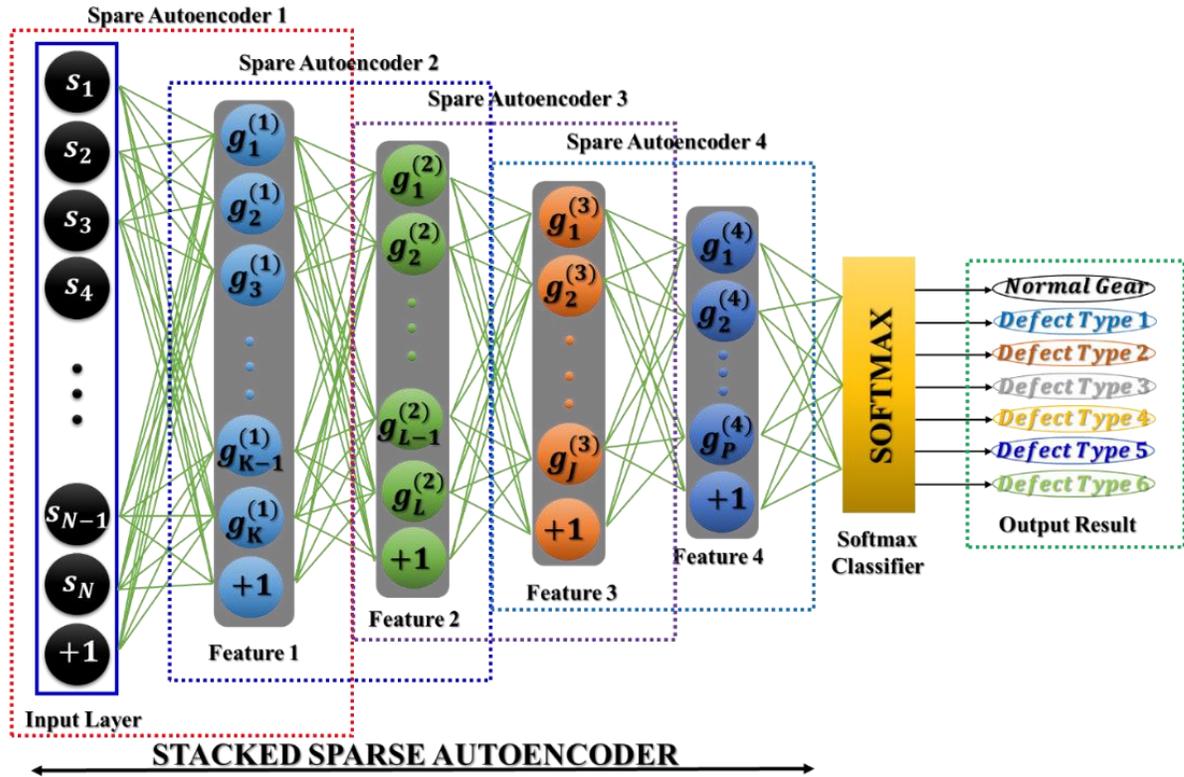
Figure 4.8. The DNA of a Stacked Sparse Autoencoder.

hidden layers, and a SoftMax classifier [140] represents an output layer of this network architecture. Hence, all sparse autoencoders, which are stacked, form the DNA. Figure 4.8 depicts an example of a DNA with four hidden layers for visual understanding. This DNA first operates in an unsupervised learning manner, where all of the SSAs extract useful features and then, in a supervised learning manner, the DNA executes fine-tuning employing a back-propagation algorithm based on the stochastic gradient descent [74]. After the training process is completed, the unseen data is used for evaluating the performance of the DNA.

4.4 Experimental Setup and Tuning DNA Parameters

To validate the effectiveness of feature engineering and classification by the SSA-DNN in the

proposed model, we perform a set of four experiments listed in Table 4.2. In these experiments, the SSA-DNN uses the input data as the samples of the frequency spectrum of the subbands that were optimized by the ANC. The four subsets of gearbox data were taken based on shaft rotation speed, i.e., each data subset contains 1400 samples in total for all defect states (200 samples for each class of seven defect states: N, D1, ..., D6), which were acquired from the vibration sensor when the shaft rotates at the same speed. For each experiment trial, the proposed DNA was trained numerous times with diverse numbers of epochs using samples corresponding to one speed of the shaft and validated with the dataset collected under two other shaft speeds, then changing samples belonging to different speeds for all four experiments.

Table 4.2. Description of the dataset for training and testing with RPM in the experiment setup.

The Experiments	Number of Samples	The RPM of Data Samples
Experiment 1	Training sample: 1400	The shaft speed: 300 RPM
	Testing samples: 2800	The shaft speed: 600 RPM and 900 RPM
Experiment 2	Training sample: 1400	The shaft speed: 600 RPM
	Testing samples: 2800	The shaft speed: 900 RPM and 1200 RPM
Experiment 3	Training sample: 1400	The shaft speed: 900 RPM
	Testing samples: 2800	The shaft speed: 300 RPM and 1200 RPM
Experiment 4	Training sample: 1400	The shaft speed: 1200 RPM
	Testing samples: 2800	The shaft speed: 300 RPM and 600 RPM

4.4.1 Tuning Parameters for the SSA-DNN

The parameters of the DNA play an important role in classification performance, so that the tuning process for selecting the optimal values has to be performed [141]. To construct this model, we have repeatedly tested the proposed model using various values of model parameters such as the length of recipient input, the sparsity regularization term, the number of hidden layers, the number of hidden nodes, and the cost function to evaluate their effect on DNA performance. The following subsections explain the parameter tuning process in detail.

The length of the recipient input is the size of a single sample which is inputted to the DNA, it is also known as the value of higher-dimensional representation of the input layer. According to [41], this parameter is the first important factor for recognizing the complex

features that can be well supported for the classification of MTCG fault types to build up the sensitive gearbox fault diagnosis model. Therefore, a larger recipient input length helps the DNA to extract better representative features. Nevertheless, a huge size of the input increases the computational complexity of the model, while a reasonable size of the input can provide both a reasonable quality of feature extraction and well-proportioned computation complexity. As mentioned in Section 4.3.1, the one second raw vibration signals were sampled at a frequency of 65,536 Hz, resulting in 65,536 points in the time do-main. This raw signal was preprocessed by three-time down sampling accompanied by low-pass filtering before entering the ANC module. Hence, there are 21,845 (65,536/3) data points in the optimized time-domain signals received in the output of the ANC module. By applying the Fourier transform to these signals, the symmetrical frequency spectrum of each optimized subband containing an imaginary part (this part represents a spectrum of the signal in the negative frequency) and a real part (for the frequency tones greater than zero) is received. The real part that represents a real frequency spectrum of an optimized subband with 10,922 (21,845/2) data points is used as the input to the DNA. The usage of a large number of data points at the input layer might increase the computational complexity; however, the effectiveness of fault identification might not be improved significantly. On the contrary, a further reduction of the input size will lead to the reduction of frequency resolution and hence, it might cause challenges for the model when identifying the MTCG defect types. Thus, the length of the recipient input with 10,922 points of an optimized subband represents a rational trade-off between the classification performance and computational complexity for the sensitive and speed invariant MTCG gearbox fault diagnosis model.

Similarly, the number of neurons in the hidden layers also influences the performance of the DNA. Although there are no exact guidelines for selecting the number of neurons for a hidden layer of an autoencoder, this parameter directly impacts the process of feature extraction. Based on the functionality of the autoencoder, the number of nodes in the first hidden layer has to be lesser than the length of the recipient input for compressing the higher-dimensional data. To adjust the parameters of node number and sparsity, in this paper we create a fine-tuning dataset which is formed by randomly picking 100 data instances corresponding to each class under each rotation speed condition. Hence, the fine-tuning dataset consisted of $N_{samp} \times N_{class} \times N_{speed} = 100 \times 7 \times 4 = 2800$ data instances in total. Figure 4.9 illustrates the relationship between the reconstruction error curve and the number of nodes for the first hidden layer obtained while training the autoencoder on the fine-tuning dataset during 350 epochs. This curve demonstrates that the number of 3000 nodes in the hidden layer, which is greater than 20% of the input size (10,922), leads to smaller reconstruction errors. A further increase in this number minorly affects the reconstruction error, but the computational
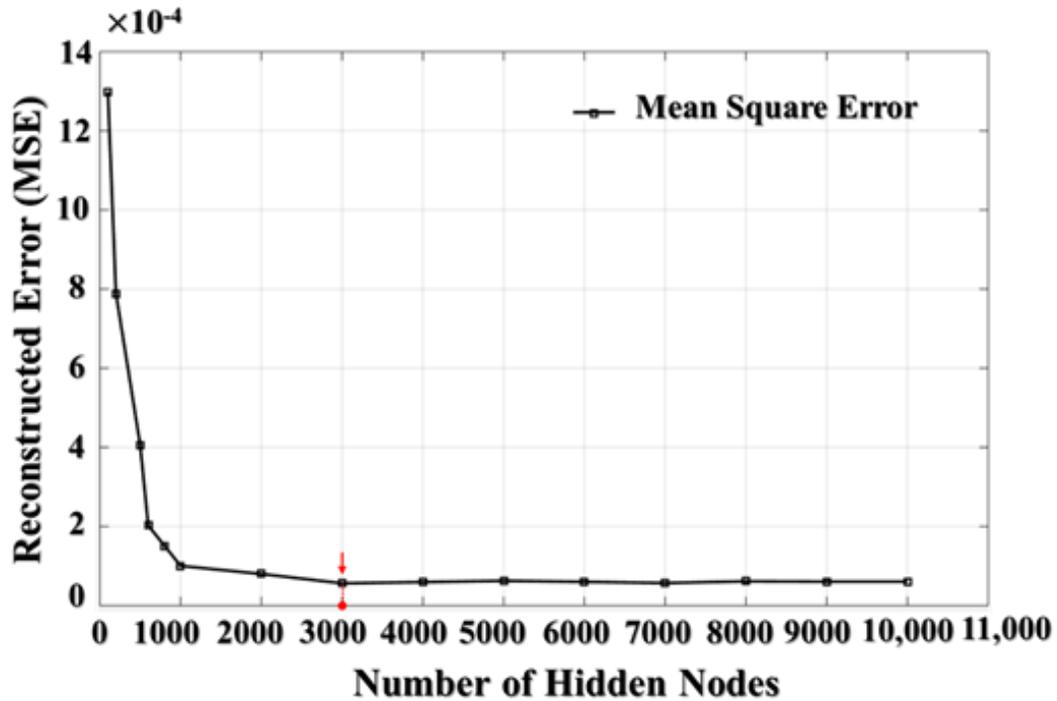
Figure 4.9. The dependence graph of reconstructed error MSE and the number of nodes in the first hidden layer.

complexity would be increased significantly. Thus, it is recommended to keep the number of nodes for the hidden layer at less than 35% of the input size. This criterion is applied to the

remaining hidden layers in the proposed model, so the number of nodes in each consecutive hidden layer is in the range from 20% to 35% of the number of nodes in the previous layer.

The sparsity penalty can be used for improving the forward learning process of an unsupervised autoencoder, whose purposive activity orients to manifest the highly representative features. To evaluate the effect of the sparsity penalty, the reconstruction error is mostly considered for the experiment the value of sparsity penalty parameter in the first autoencoder (the first hidden layer is selected with number hidden nodes as 3000). Figure 4.10 demonstrates the relation between the value of the sparsity term and the reconstruction error, which is a mean square error (MSE) in this study, achieved when training the auto-encoder on a fine-tuning dataset during 350 epochs. It is observed that values of sparsity penalty in the range from 0.05 to 0.15 are better than the remaining values, and a value of 0.08 is the optimal one leading to the minimum MSE. Hence, this value has been chosen as a penalty factor for
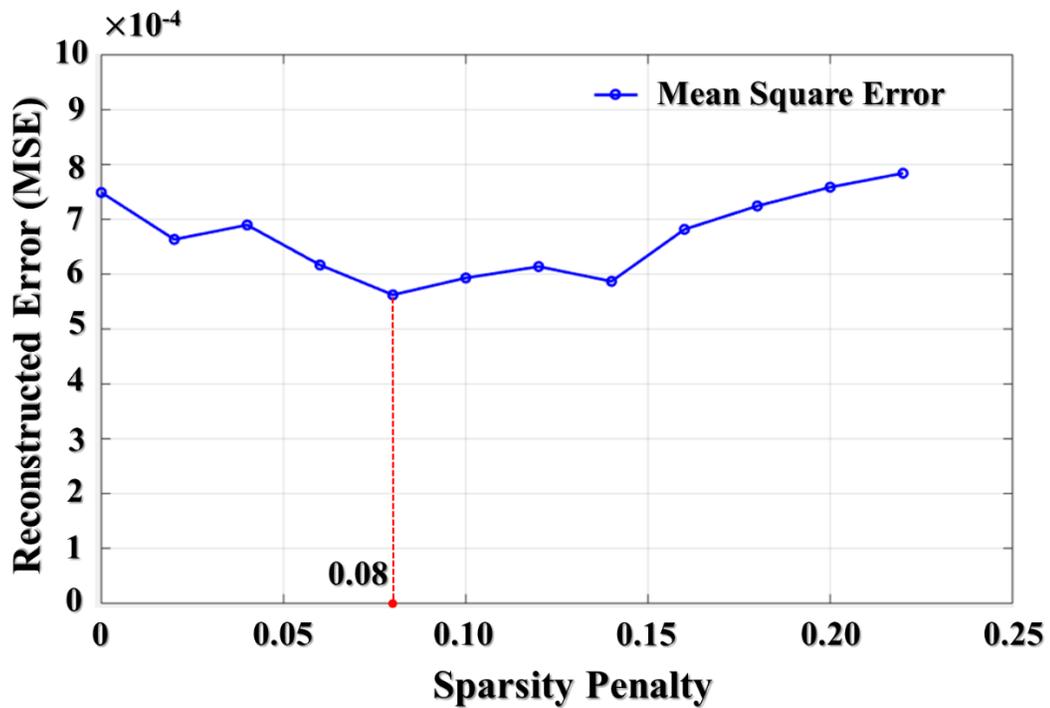
Figure 4.10. The relation graph between the sparsity term and the reconstruction MSE.

all the hidden layers in the proposed model.

The number of hidden layers plays an important role in the learning process. There exists a general opinion that a higher number of hidden layers results in better accuracy, but also

reduces the generalization ability of the network [74]. In this work, a series of experiments to determine the number of hidden layers were performed while varying their number from three to six, as shown in Table 4.3. From this table, it can be observed that a number of hidden layers greater than three leads to the smallest reconstruction errors. Regarding a higher number of hidden layers, the reconstruction error does not change significantly; however, the computational time can be increased dramatically when making the architecture deeper. Therefore, to select a suitable number of hidden layers, the time performance also should be considered.

Table 4.3. The reconstruction error with the sets of numbers of hidden layers and their nodes.

Number of Hidden Layers	Nodes per Each Layer	Reconstruction Error
3	3500/1500/500	16.312×10^{-3}
3	3000/1000/300	15.189×10^{-3}
4	4000/1800/600/200	9.745×10^{-5}
4	3500/1500/500/200	6.887×10^{-5}
4	3000/1000/300/100	4.698×10^{-5}
5	6000/2000/800/250/80	3.783×10^{-5}
5	5000/1800/600/200/60	4.2×10^{-5}
5	4000/1400/400/160/60	4.034×10^{-5}
6	8000/4000/1500/500/200/60	1.439×10^{-5}
6	7000/3000/1000/400/150/50	1.907×10^{-5}
6	6000/2500/800/300/100/50	2.543×10^{-5}

The complexity of computation of the architecture, in general, can be measured as an average time required for one training cycle of DNA. Figure 4.11 shows the time consumption of different SSA-DNN deep architectures with various numbers of hidden layers and nodes in them during the training process. In this figure, the DNAs with higher numbers of hidden

layers and nodes requires more time for training due to the depth of the architecture.

4.4.2 Parameter Selections of the SSA-DNN Model
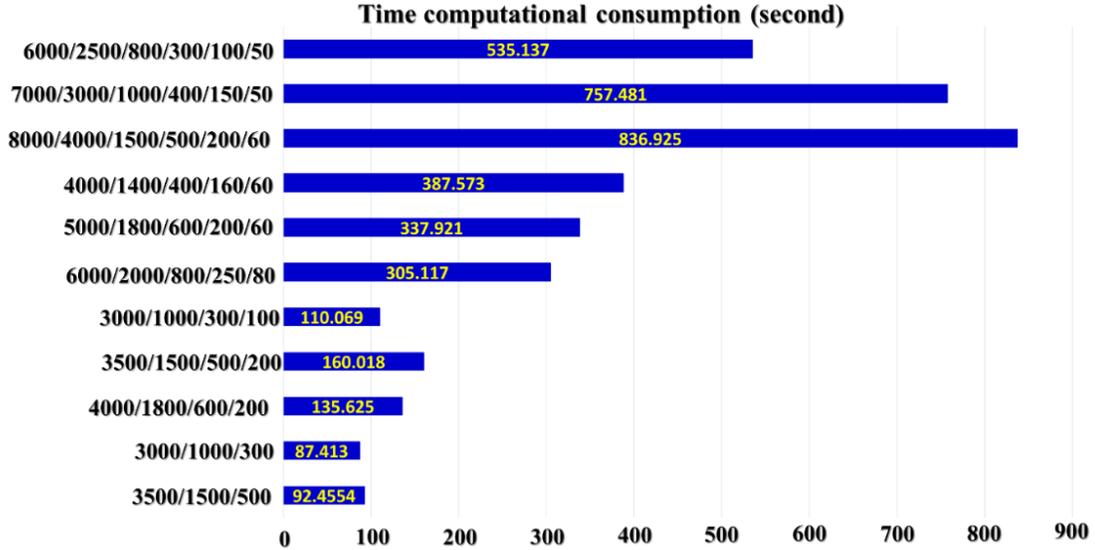
Figure 4.11. The training time consumption of SSA-DNN architectures with different numbers of hidden layers and nodes in them.

Through the experiments in the previous subsection, it was observed that with the increase of DNA architecture complexity, the reconstruction error was getting smaller while the time needed for training the deep architecture was increasing. However, from Table 4.3 it can be seen that after reaching certain numbers of hidden layers and nodes, the further increase of architecture complexity leads only to minor reductions of the reconstruction error. From this observation, it can be concluded that the actual number of highly representative features is limited, and thus, when the DNA attempts to extract more features from its input, which might be redundant and not representative, they would not affect the resulting reconstruction error significantly. The structure of a DNA should contain several numbers of hidden layers to adequately perform dimensionality reduction of the input data, where each hidden layer analyzes its input to perform both feature extraction and selection to receive the higher-level representative features. These features are then used for discriminating the MTCG defect types during the classification process. Because of the challenge of constructing the speed invariant fault diagnosis model for MTCG gearbox systems, the parameters are selected to prioritize the small reconstruction error with acceptable execution time consumption.

Regarding the architectures with five or six hidden layers, the reconstruction errors are relatively small in comparison with other architectures, though, the time consumed for the training process is much higher and the error values are not much larger. Therefore, in this study, the number of hidden layers is selected as four with the amounts of nodes (i.e., number of features) in them as 3000, 1000, 300, and 100 neurons for the first, second, third, and fourth hidden layers, respectively. The finalized optimal parameters of the SSA-DNN model are listed in Table 4.4, and its architecture is shown in Figure 4.12.

Table 4.4. The optimal selected parameters for constructing the SSA-DNN model.

Input Size (Sample Length)	Number of Layers	Number of Nodes	Sparsity Constraint	Activation Function
10,922	4	3000, 1000, 300, 100	0.08, 0.08, 0.08, 0.08	Logistic sigmoid
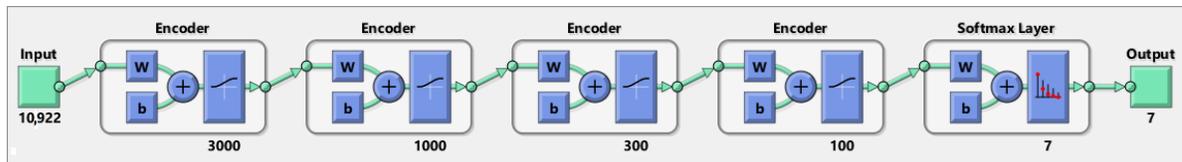
Figure 4.12. The final architecture of the SSA-DNN model.

4.5 Result and Discussion

The main function of the ANC is to perform noise reduction and to preserve the fault-related useful components existing in the vibration signals. To collect the informative content of the vibration sample, where the content represents numerous fault-related components that are useful for designing the sensitive fault diagnosis model, the analog signals from the vibration acceleration sensor were digitized with a high sampling frequency of 65,536 Hz every one second. Thus, a 1-sec length data sample is used to monitor several rotation cycles (from three to thirteen rotational cycles depending on the rotation speed from 300 RPM to 1200 RPM) to collect fault-related vibration characteristics with some special oscillations. After data collection, the digitized vibration samples were filtered by a digital low-pass filter with the cut-off frequency of 10,000 Hz accompanied with the down-sampling process to remove the high-frequency components (i.e., components located in spectrum higher than 10,000 Hz) which are out of operation range of the acceleration sensor, and to preserve the vibration

components with intrinsic fault-informative features following realistic operation of a gearbox system. That is the first step for preprocessing data to remove the redundancy in the raw vibration signals. The vibration subbands output from a low-pass filter are inputted into the ANC module for a fine-optimizing process for noise reduction. In the range of the frequency spectrum less than 10 kHz, the ANC uses adaptive windows to access and remove white noise and band noise remaining between two consecutive sideband frequencies along the frequency spectrum.

Figure 4.13 demonstrates the superiority of the ANC module for the de-noising process. Here, the red dotted circles indicate the noise frequency component zones of the input signals which were reduced significantly in the optimized subband outputted from the ANC. Moreover, the amplitudes of the sideband frequency tones, the meshing frequency, and its harmonics are kept unchanged when the vibration subband flows through the ANC module (the dashed blue and black circles). The outputs of the ANC are the optimized vibration subbands represented in the frequency domain for the expression of the energy distribution. These spectra are used as inputs to the SSA-DNN module for extracting the representative latent features by an

unsupervised learning technique, the autoencoder, which is a part of the SSA-DNN module.
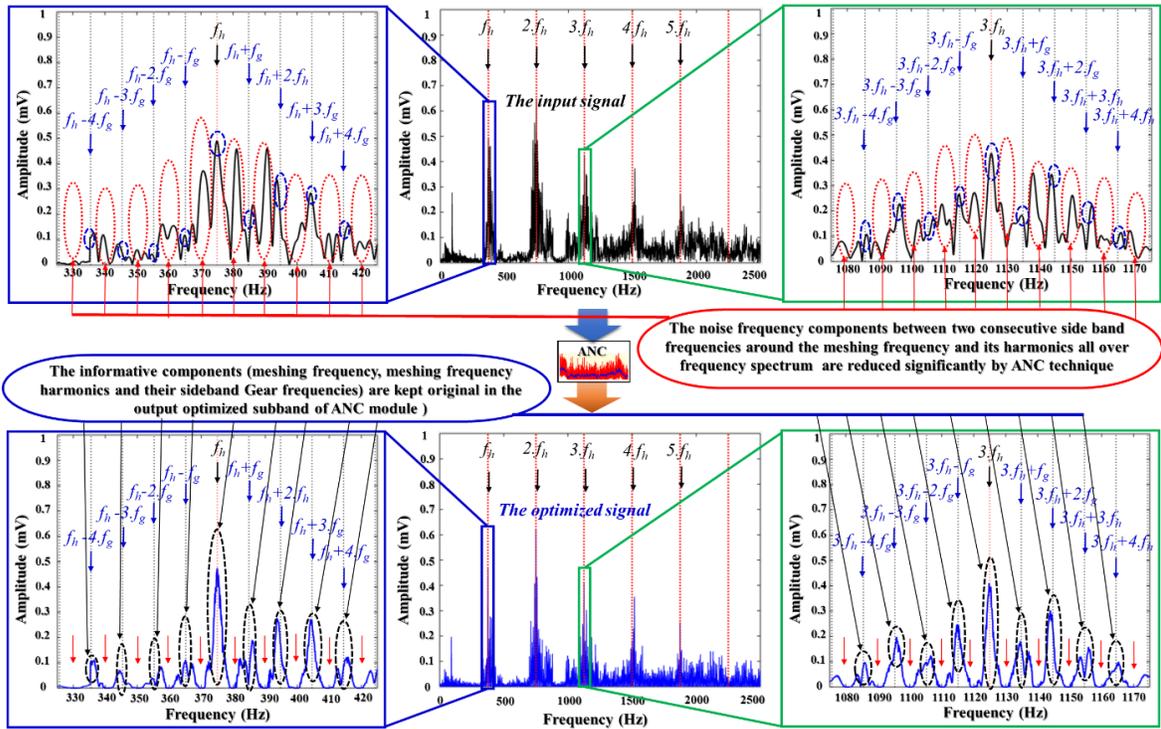
Figure 4.13. Frequency spectrum analysis of the vibration subband (for fault state D2 at 900 RPM) in the comparison between an input and output subband of the ANC module.

sparse autoencoders under four various rotational speeds.

Figure 4.14 illustrates feature spaces for seven defect types of an MTCG gearbox using some of the discriminative features extracted by sparse autoencoders from the frequency spectra of optimized subbands under different rotational speeds. This figure shows that the data instances corresponding to different signal classes are well separable in feature space. Here, the samples belonging to one defect type are placed closely, whereas the samples of different defect types are located separately in the visualized feature space. These distinct features are extracted by stacking the sparse autoencoder layers and are used to enhance the performance of the deep architecture using a back-propagation algorithm to minimize the reconstruction errors and then, finally, to classify gearbox defects. For fault diagnosis performance evaluation, we compared the results of the proposed model with previous models such as ANC and SVM [126] (model 1), ANC and ANN (model 2), stacked denoising autoencoder [142] (model 3), and the spectra imaging of vibration signal [143] (model 4).

These results are presented in Table 4.5. The performance is evaluated using the four cases of experiment setup expressed in Table 2. The training dataset of each experiment contains 1400 vibration samples (200 vibration samples for each defect state of seven states as N, D1, D2, D3, D4, D5, D6) for each rotational speed to construct the deep architecture network model. The testing process is performed by 2800 vibration samples of two different rotational speeds. By executing four experiments, the vibration samples of four rotational speeds are used for training set in sequence, whereas two datasets of rotational speeds, which are different from rotational speed in training dataset in each experiment, are consumed for the testing process. In these experiments, models 1 and 2 use the statistical features extracted from time and frequency domains, whereas the remaining models use autonomous feature extraction methods based on the unsupervised learning approach (model 3) and vibration imaging approach (model 4). Models 1 and 2 use the optimized subband output from an ANC module to extract twenty-one feature parameters and then, using these feature vectors, classify fault types using SVM and ANN, respectively. Manually extracted features in models 1 and 2 cause a challenge when classifying multi-level tooth cut gear defects. Their fault classification results were around 68%±10% for model 1 and 59.4% ± 10% for model 2, fluctuating over four experiments. The construction of DNA in model 3 is performed by replacing the four sparsity autoencoder hidden layers with two layers of denoising autoencoders using the optimal regularization terms and parameters from [142] and removing the ANC module from the proposed model. In model 3, the input data are the vibration subbands outputted from the down sampling and low-pass filtering process, with the denoising and feature engineering processes performed using the objective functions with the embedded manifold regularization. The fault identification results achieved by this model were about 82.88%±8% in four experiments.

Table 4.5. Classification results of the referenced and proposed models in four experiments based on various rotating speed data.

Models	Training Set (1400 Samples)	Test Set (2800 Samples)	Accuracy
	300 RPM	600RPM, 900 RPM	62.78
	600 RPM	900RPM, 1200 RPM	79.83
I	900 RPM	300RPM, 1200 RPM	67.13
	1200 RPM	300RPM, 600 RPM	62
	Average accuracy by four experiments		**68**
II	300 RPM	600RPM, 900 RPM	65

	600 RPM	900RPM, 1200 RPM	48.10
	900 RPM	300RPM, 1200 RPM	73.5
	1200 RPM	300RPM, 600 RPM	51
	Average accuracy by four experiments		**59.4**
	300 RPM	600RPM, 900 RPM	90.66
	600 RPM	900RPM, 1200 RPM	79
III	900 RPM	300RPM, 1200 RPM	85.50
	1200 RPM	300RPM, 600 RPM	76.35
	Average accuracy by four experiments		**82.88**
	300 RPM	600RPM, 900 RPM	41.15
	600 RPM	900RPM, 1200 RPM	39.55
IV	900 RPM	300RPM, 1200 RPM	48.26
	1200 RPM	300RPM, 600 RPM	51.72
	Average accuracy by four experiments		**45.17**
	300 RPM	600RPM, 900 RPM	95.51
	600 RPM	900RPM, 1200 RPM	97.32
The proposed model	900 RPM	300RPM, 1200 RPM	99
	1200 RPM	300RPM, 600 RPM	96.1
	Average accuracy by four experiments		**97**

These results can be observed because many fault-related components stay hidden in the background noise which can only be detected by the application of signal processing methods. Regarding model 4, the raw 1-sec vibration signal with 65,536 points was firstly down-sampled by four times with a 10 kHz low-pass filter integrated for antialiasing to obtain the vibration subband with 16,384 data points. Then this subband is segmented in series without overlap by using windows of 1024-point size to attain sixteen segments of 1024-point vibration subbands. Then, each 1024-point window containing the vibration subband is transformed from the time domain to the frequency domain by FFT to obtain a 513-point sized vibration frequency spectrum. This process was repeated eight times by randomly picking eight segments of 1024-point vibration subbands from sixteen segments. These spectrums were stacked to form the 513×8 grayscale image corresponding to each raw 1-sec vibration sample. This image was later converted to a binary image by an 8×4 sized filter and the threshold (0.7). Hence, the binary image containing 4014 components in the frequency domain was used as the input to the ANN with three layers (input, hidden with three nodes, and output layers) for classification. The fault classification results of model 4 on the dataset used in this paper were about 45.17% $\pm$ 6% during four experiments. By analyzing the experimental results of the referenced models, it can be seen that the sensitive and speed

invariant fault diagnosis model proposed in this study outperformed their fault diagnosis performance with results around 97% ± 2% during four experiments showing small accuracy deviations when alternating the shaft rotational speeds of the MTCG gearbox system. Additionally, to verify the stability of the proposed algorithm, the experiments de-scribed above have been performed five times. The classification accuracies and their averages computed over five experimental trials are presented in Table 4.6. From these results, it can be seen that the proposed model demonstrates stable fault classification accuracy in independent trials of the experiments performed for training and testing subsets containing samples collected under different operating conditions, i.e., rotating speed.

Table 4. 6. Fault classification results of the proposed model obtained during five experimental trials.

Training Set (1400 Samples)	Testing Set (2800 Samples)	Experiment Trials					Average Accuracy (%)
		#1	#2	#3	#4	#5	
300 RPM	600 RPM 900 RPM	93	96.87	95.7	93.85	98.15	95.51
600 RPM	900 RPM 1200 RPM	97.68	98.2	94.95	100	95.78	97.32
900 RPM	300 RPM 1200 RPM	100	100	99.68	98.19	97.15	99.00
1200 RPM	300 RPM 600 RPM	98.00	94.28	95.47	97.9	94.87	96.10

Controlling the noise embedded in the vibration signals is essential for the sensitive detection of multi-level cut tooth faults in gearbox systems. The presence of a high noise level can cause misidentifications of fault types and thus reduce the fault classification accuracy. Noise reduction is a complex problem, and it is not always possible to completely resolve this issue by signal processing or feature engineering techniques. Therefore, simultaneous usage of the ANC and SSA-DNN methods is an efficient approach for significant noise reduction while preserving the original fault-related information of the gear vibration characteristic, which is useful for fault identification. The design of a sensitive and speed invariant model requires exploration of the representative features that can be used for discrimination of multi-level tooth cut gear defects and maintaining its reliable performance under the operating speed fluctuation conditions in the gearbox system. In general, the manual feature extraction methods cannot satisfy those requirements, thus the unsupervised approaches based on deep neural networks are well-suitable for extracting the latent representative features by the

process of minimizing reconstruction errors during the operation of a back-propagation algorithm in the DNA. The SSA-based DNN constructed in this research satisfies the requirements for constructing the proposed model, such as extracting the representative feature space, selecting the most defect-related useful features for classification, and finally, achieving high fault classification results.

4.6 Conclusion

This study presents a novel method which combines an ANC and an SSA-DNN to utilize their advantages for constructing a sensitive and speed invariant fault identification model for gearbox systems with multi-level tooth cut gear defects. The ANC technique is created based on the analysis of vibration characteristics of a gearbox system to generate the speed-dependent reference window signals with adjustable parameters, according to the noise types presenting in the raw vibration signals. Then, these generated window series were adaptively adjusted to access the space between two consecutive defect-related frequency tones and remove the noise along the whole frequency range of vibration signals. The ANC optimizes the input vibration signal for outputting the optimal subband which contains mostly the defect-related frequency tones with the integration of low-level background noise. Then, the frequency spectra of these optimal subbands are used as the input to the deep network architecture. This network is built up by stacking sparse autoencoders as the hidden layers of the network and using a Softmax activation function at the output layer for extracting latent representative feature spaces and selecting the most defect-related discriminative features for identifying the multi-level tooth cut fault types under the condition of various shaft rotational speeds. The effectiveness of the proposed model is validated by experiments performed using the vibration dataset containing MTCG gearbox defects collected under four different rotational speeds. To validate the property of speed invariance for the proposed model, the experiment was arranged as four sub-experiments using the datasets corresponding to each rotational speed. Each sub-experiment uses a one-speed dataset to construct and train the model. Then this given model is used for fault identification using two datasets collected under other speed conditions. This procedure was performed four times using the different speed datasets for building the model in each. The average classification result achieved over four experiments was 97%, which outperforms the techniques used for comparison. Moreover, the classification results shown by the proposed model did not fluctuate significantly (2–3%) when applied to different speed datasets, which evidences that the prosed model is speed invariant and can be used for identifying multi-level tooth cut defects in a gearbox system under varying rotational speeds.

Chapter 5

Gearbox Fault Identification Framework Based on Novel Localized Adaptive Denoising Technique, Wavelet-based Vibration Imaging, and Deep Convolutional Neural Network

5.1 Introduction

Gearboxes play an important role in numerous industrial machines, vehicles, and wind turbines [27][144][145]. Due to the operation of gearboxes in harsh conditions, gear defects are found to be the most common defects in gearboxes [5]. A fault in the gearbox can result in catastrophic failures, economic losses, and danger to the operating staff. For this reason, early fault detection of the gearbox is of primary importance. The condition based monitoring approach suggests maintenance action based on the data collected from the gearbox. This strategy allows the gearbox to function for a long time with minimal maintenance costs [146][14].

Gear fault signatures are sensed and acquired by two types of sensors: accelerometers and acoustic emission sensors [83]. Vibration signatures collected by the accelerometer from a gearbox carry enough fault-related information and can be used for efficient gear fault diagnosis [84]. Vibration signals obtained from a gearbox consist of meshing frequency harmonics, blended sideband frequencies, and other free oscillation frequencies. Therein, the meshing frequency harmonics and blended sideband frequencies are the fundamental defect-related frequencies that help in the process of identifying gear defects [3][80]. The vibration signals obtained from a gearbox under variable speeds are complex and nonstationary; furthermore, the gear fault-related elements are often overwhelmed by the

noise. To identify the fault symptoms in this complex vibration signal, the fault diagnosis technique tries to reduce the noise in the raw vibration signal [132]. In its raw form, the gearbox vibration signal contains various types of interference noises. The main sources for these interference noises are the interconnected systems, such as the electrical-electronic control and measuring systems, the mechanical systems (the influence of the mechanical resonances such as shaft, bearings, gears, etc.), and background noise [147][85]. The random behavior of these noises (i.e., random magnitudes, random appearances anywhere in the observed ranges of vibration signals) makes the noisy components dominant over the fault-related components in the vibration signal, and thus these noises overwhelm the fault-related components. To address this issue, a signal-processing technique, which can reduce the noise in the raw vibration signal, to highlight the fault-related meshing frequency harmonics and sidebands (fault-affiliated elements) for gearbox fault diagnosis in early stages is urgently needed.

In the past, numerous signal processing techniques, such as Fourier transform (FT), envelope spectral analysis, Hilbert transform (HT), spectrogram or spectral analysis of a fixed timing-window Fourier transform (STHT), empirical mode decomposition (EMD), and wavelet-based spectral analysis (WA), have been developed for the processing of stationary and non-stationary complex signals [91][87][100][101][102][99]. To enhance the performance of the basic signal processing techniques, hybrid signal processing techniques such as EMD and HT and EMD and WA have also been introduced [105][53]. The vibration signatures obtained from a gearbox under faulty conditions are non-stationary. To obtain fault-related information from the non-stationary vibration signal, time–frequency domain techniques are applied. These techniques commonly use window-based filtering, digital filtering, threshold estimation, decomposition modes in the form of intrinsic mode functions, and wavelet-based transformation. Their fault identification efficiencies have been confirmed by classifying fault states (e.g., a fault-free state and a defective state) and denoising in some cases. In the case of a gearbox, the fault-related information is distorted by the hefty noise present in the raw signal. Therefore, noise reduction techniques before applying the time–frequency domain signal processing technique will be helpful for the identification of MDTF in gearboxes. Nguyen et al. [126][148] proposed an adaptive noise reduction model, which effectively reduced the noise in the impaired signal. The resultant impaired signal

is then used for the classification of gearbox multi-level tooth cut faults under variable speed conditions. The effectiveness of the adaptive noise reduction model lies in adaptively adjusting optimal parameters of the Gaussian function, which are connected to the optimal weights of the adaptive filter, along the whole frequency range of a vibration signal. Nevertheless, the frequency spectrum of a vibration signal obtained from a gearbox is composed of meshing frequency harmonics, sideband components, and random noises, with different probability distributions. It should be noted that the influence of random noises and the change in stiffness of the gear under defect makes the vibration signal nonstationary and complex. Therefore, a single optimal parameter set of the Gaussian reference signal along the entire frequency range is less effective for noise reduction. To address this issue, a localized adaptive denoising technique (LADT) is proposed in this paper. The proposed LADT is a modified version of the adaptive noise reduction model proposed in [126]. The LADT adaptively transforms the raw vibration signals to the optimized subbands, which accounts for the majority of the defect-related information. The proposed method can reduce noise more effectively than the previous adaptive denoising models, while maintaining original fault-related information. The resultant impaired signal from LADT is then used for feature engineering and fault classification in the proposed scheme of gearbox fault diagnosis.

After signal preprocessing, feature preprocessing and fault classification are the most important steps in the fault identification system. Conventional methods for the fault diagnosis of gearboxes used handcrafted features. After extracting a limited number of features from the signal in the conventional methods, domain knowledge was used for discriminant feature selection. These discriminant features were then classified using machine learning algorithms, such as support vector machines (SVMs), k-nearest neighbors (KNNs), decision tree algorithms (DTAs), and artificial neural networks (ANNs) [66][67][133][149][150]. However, the handcrafted features need domain knowledge and expertise for the identification of discriminant features. Furthermore, feature engineering techniques, such as dimensionality reduction for discriminant feature selection, result in fault-related information loss. Thus, the conventional methods might not be appropriate for the classification and identification of MDTF defects in the gearbox. In addition, classification algorithms, such as KNNs, SVMs, DTAs, and ANNs, is strongly dependent on the quality of

the provided features. To address the above-discussed problems, this paper proposes a scheme of self-generating feature space. The proposed scheme first transforms a low-noise vibration signal into a two-dimensional (2D) image using wavelet transform and obtains WVI's. The WVI's reflect the 2D distributed power spectra of the optimized vibration sub-bands. To obtain fault-related information from the WVI and classify them into their representative classes, the proposed method used DCNA. Deep learning models (DLMs) have been used widely in the areas of finance, natural language processing, and image processing [151][152][45][77]. For condition monitoring of a rotating machine, there exist a variety of DLMs based fault diagnosis frameworks, such as stacked denoising autoencoder [153], recurrent neural network [154], long short term memory (LSTM) networks [155], gated recurrent unit network [156], and convolutional neural network (CNN) [156][157]. One of the deep learning models, CNN, is a famous model because of its visual understanding [72]. Deep convolutional network architecture (DCNA) has been created for image processing and recognition, and then developed for fault diagnosis of rotation types of machinery by self-regulation and deep exploration of the latent fault-reflected features of vibration signals [78], [158]–[160].

The contributions of this study are briefly explained as follows:

(1) A new signal preprocessing approach LADT is developed. The LADT is an adaptive algorithm that considers each principal frequency segment along the frequency spectrum of a vibration signal to fetch the optimized Gaussian parameters, called localized optimal parameters. The outputs of the LADT, which are optimized vibration sub-bands, contain fault-related information with very low interference noise.

(2) To discriminate and highlight the fault-related information in the vibration signals of MDTF defect types in the time–frequency domain, the WVI technique is applied.

(3) Potential features are extracted from the WVI's and classified using DCNA. The latent features of DCNA contain discriminant fault-related features. To classify the fault-related features into their respective classes, DCNA then uses the fine-tuning process based on the backpropagation algorithm.

he remaining sections of this work are arranged as follows: the vibration characteristics of the gearbox are explained in Section 2; Section 3 describes the technical background. The experimental setup and proposed diagnosis scheme are explained in Section 4. Section 5 presents the discussion and evaluation of the experimental results obtained from the proposed scheme, and finally, the conclusion of this study is presented in Section 6.

5.2 The Specification of a Gearbox Vibration Signal

A fault in the gear results in a change in the stiffness. This stiffness can be observed in the vibration spectrum at specific characteristic frequencies. These characteristic frequencies represent the tooth meshing stiffness. The meshing frequency in the vibration spectrum of the gearbox represents the symptoms of a defect in the gearbox, as the meshing frequency changes whenever an MDTF occurs in the gearbox [161]. Considering a gearbox operating under normal conditions, the vibration signature obtained from the gearbox is a stationary signal with tooth meshing frequency; this signal can be formulated as follows [81].

$$x_g(t) = \sum_{p=i}^{P} X_p \cos(2\pi p f_m t + \varepsilon_p), \tag{5.1}$$

where $x_g(t)$ represents vibration signal of a gear operating under normal condition, X_p and ε_p stand for amplitude and phase of p-th harmonic of a meshing frequency, P denotes harmonics of the meshing frequency; and f_m, which denotes meshing frequency, it can be computed using the parameter of pinion wheel (f_m = Number of pinion teeth × Rotational frequency of a pinion wheel) or using a gear wheel (f_m = Number of gear teeth × Rotational frequency of a gear wheel). The meshing frequency and its harmonics are considered useful components for the fault diagnosis process. Figure 5.1a shows an example of the frequency spectrum of a vibration signal in the perfect condition.

A fault in the gearbox makes the vibration signal non-stationary, resulting in a complex frequency spectrum. During the gearbox operation, transmission occurs between the motion source (e.g., three-phase motor and a drive shaft) and a load (a non-drive-shaft and a load) through a pair of gears (pinion wheel and gear wheel). The non-stationary impulses start appearing in the vibration signal when there is an impulsive change in the angular velocity. The angular velocity changes impulsively when the two wheels

rotate across a faulty tooth (e.g., missing tooth, cracked tooth, chipped tooth, or worn tooth) [82]. Therefore, the vibration signals obtained from a faulty gearbox exhibit nonstationary behavior, for which the frequency spectrum contains harmonics of tooth meshing frequency, sidebands (the frequency tones are distributed in the two sides of harmonics of a meshing frequency), and other oscillation components. The vibration signal can be presented as a combination of phase and amplitude modulation signal [121], as follows:

$$x_m(t) = \sum_{p=0}^{P} X_p \left(1 + \beta_p(t)\right) \cos\left(2\pi p f_m t + \varphi_p + \phi_p(t)\right). \tag{5.2}$$

Here, $\beta_p(t) = \sum_{q=0}^{Q} B_{pq} \cos(2\pi q f_b t + \gamma_{pq})$ and $\phi_p(t) = \sum_{q=0}^{Q} \Phi_{pq} \cos(2\pi q f_b t + \varepsilon_{pq})$ represent the amplitude and phase modulation functions of the defective vibration signal. f_b is the sideband frequency, Q stands for a total number of sideband tones around p-th harmonics, B_{pq}, Φ_{pq} represents the amplitudes and γ_{pq}, ε_{pq} denote phases of q-th sideband in the amplitude, phase modulation functions, respectively. Figure 5.1b shows the frequency spectrum of the vibration signal obtain from the gearbox under defective conditions, the fault signatures or fault-related components are the harmonics of meshing frequency and sideband frequencies.
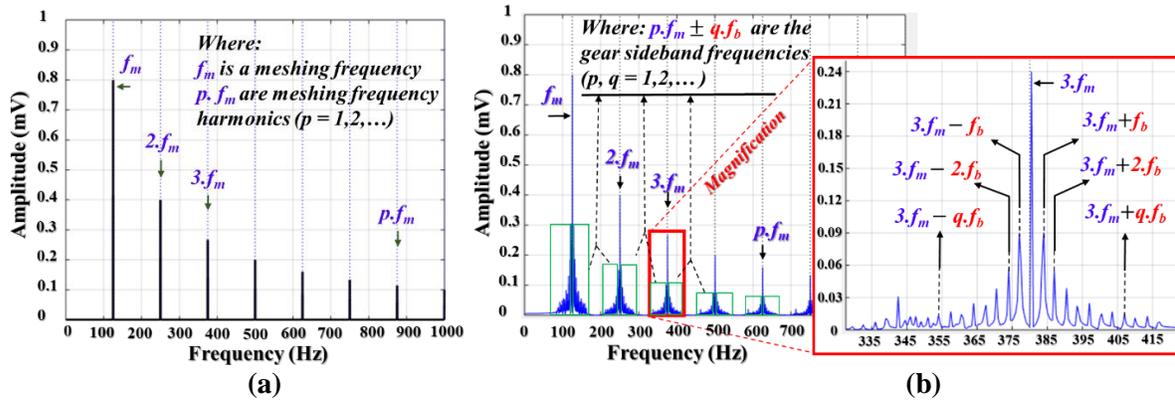
Figure 5.1. The frequency spectrum of a gearbox (a) under normal condition (b) under defective condition.

5.3 The Preliminaries

This section provides insight into the methods used in the proposed gearbox fault diagnosis scheme.

5.3.1 The Proposed Localized Adaptive Denoising Technique

Generally, a vibration signal obtained from a gearbox contains fault-related vibration signatures and noise. Denoising of the signal is required for the extraction of fault-related vibration signatures. Suppose the observed signal is s and the informative signal is x; then $s = x + \partial$, where ∂ represents the noise. The denoising technique tries to filter out noise for obtaining the estimation signal $\hat{x}$ in a tendency to approximate the useful signal x as much as possible. The adaptive denoising technique uses the concept of destructive interference for denoising of an impaired signal. This technique utilizes the noise-simulated reference signal to access frequency segments in a frequency domain of the observed impaired signal in order to remove noise. The adaptive noise-reducer-based Gaussian reference signal (ANR-GRS), which has been proposed and verified in [126][148], has achieved great performance in reducing noise and avoiding distortion of the fault-related ingredients. In this method, the noise ∂ in a gearbox vibration signal is analyzed and divided into two types of noise: white noise (α) and band noise (β), $\partial = \alpha + \beta$. Then, the reference signal is created by combining two noise-simulated signals, which are analogous with two existing sources of noise in the observed signal, a white noise signal and a Gaussian signal. Moreover, the parameters of the reference signal are adjustable by adaptive algorithm regarding the varying input values of rotation shaft speeds..

The Gaussian signal is responsible for building the simulated noise reference signal. The parameters of the Gaussian signal (a mean value and a standard deviation value) are adaptively adjusted so as to reduce the noise between two consecutive sideband frequencies (the sideband frequency is the gear frequency in this study). The process for generating a reference signal is depicted in Figure 5.2, and the Gaussian signal is formulated as follows:

$$G_{ns}(k) = \sum_{k=1}^{K} e^{-\frac{(k-F_m)^2}{2\sigma^2}} , \qquad (5.3)$$

where, K is the number of sideband segments, the mean value F_m and the standard deviation value σ are the functions of the shaft rotation frequency. Those parameters are adjusted by an optimization process to select the optimized vibration sub-band as an output of the ANR-GRS module [126].

From each parameter set (Fm, σ), which is randomly selected from the specific required range defined in [126], a noise-simulated signal is generated. This reference signal is provided as an input to the adaptive filter along with the impaired observed signal. The adaptive noise filter contains a digital filter, which employs an L-tap FIR type digital filter and weight vector as $w(n) \equiv \equiv [w_0, w_1, \ldots, w_{L-1}]^T$, and a least mean square adaptive algorithm. The adaptive filter works as follows: The noise-simulated signal is provided as an input to the digital filter, then the filtered output signal is summed with the vibration signal (impaired observed signal) to compute the error signal. This output error signal is provided as a feedback input to the adaptive algorithm to measure its mean square value. Next, the adaptive algorithm tunes the weights of the digital filter according to the converging criterion of least mean square (LMS) error to obtain the optimal weight vector (w_o) and then expose the optimal vibration sub-band corresponding to the particular parameter set. The schematic diagram of the ANR-GRS is provided in Figure 5.2.
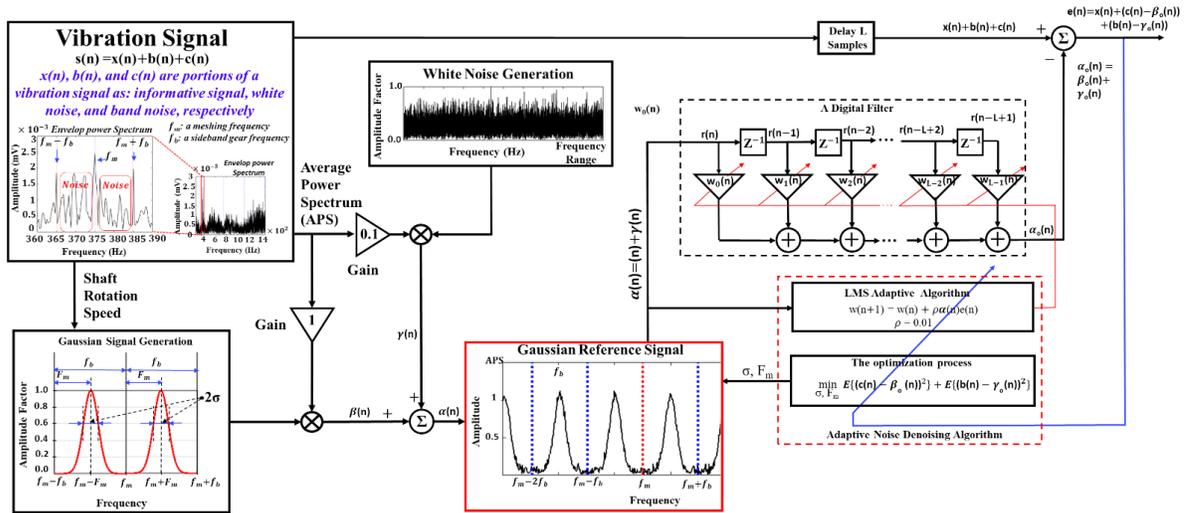
Figure 5.2. The schematic diagram of the ANR-GRS module.

From Figure 5.2, it can be observed that the ANR-GRS method tries to look for the general optimal parameters of the Gaussian reference signal applied to the whole frequency

range of input vibration signals (0–10 kHz).

According to the vibration characteristic of the fault signal presented in section 5.2, the frequency domain of the phase-amplitude modulation signals is visualized as a set of many similar frequency segments, each of which contains a meshing frequency harmonic as a center frequency and the sideband gear frequency tones are distributed around the center frequency in the ideal condition. The principal frequency segment (PFS) is defined as a frequency segment with a meshing frequency harmonic as a center frequency and frequency wide equally to a meshing frequency (i.e. the frequency range of PFS is from $(p-0.5)*f_m$ to $(p+0.5)*f_m$ with $p*f_m$, a p-th harmonic of a meshing frequency, is a center value). However, in the real world, the amplitudes of frequency tones in each PFS (PFS power distribution) of the gearbox vibration signals are uncorrelated to each other because of the influence of random noise (white-nose and band noise) on the nonlinear and phase-amplitude modulation signal [23].

Due to the differences of power distributions of PFSs, the general optimal parameter set of Gaussian reference signals cannot be used. Therefore, this paper proposes a new denoising technique called the localized adaptive denoising technique (LADT). The localized adaptive denoising technique adopts the ANR-GRS module from [126]. To improve the denoising capability of ANR-GRS, the LADT applies ANR-GRS to each PFS. By localized adaptive optimization, the new denoising methodology tries to find the localized optimal
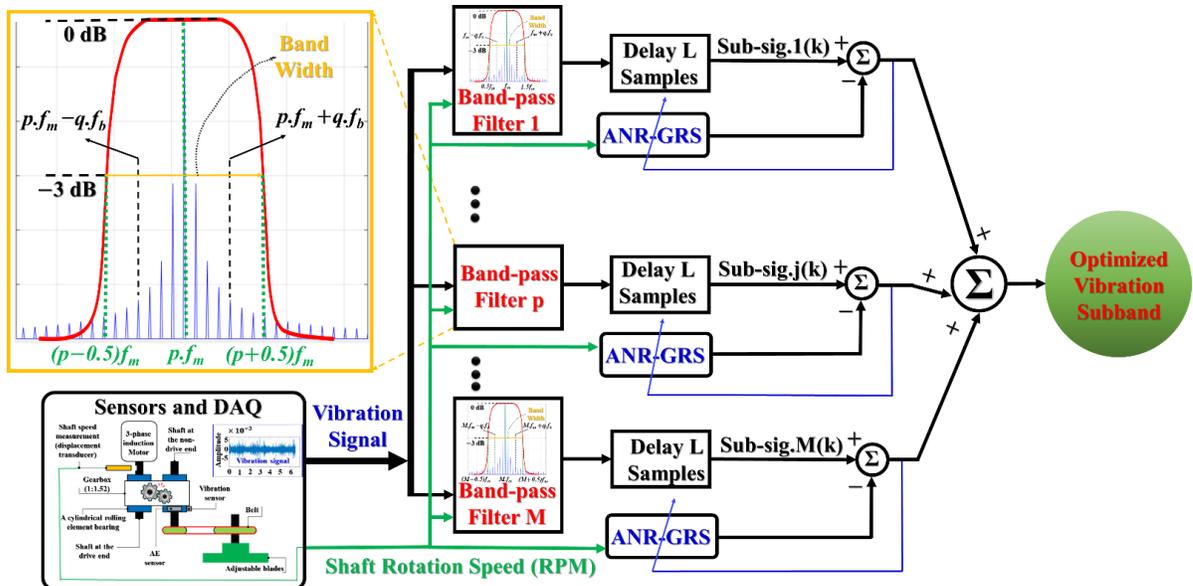
Figure 5. 3. Block diagram of the LADT.

parameter set of a noise-simulated reference signal, which is appropriate to each specific PFS. The function block diagram of LADT is demonstrated in Figure 5.3. To implement the ANR-GRS method on each PFS, the band-pass Chebyshev Type-I IIR filter of order 30 [162] is used to segment the frequency spectrums of a vibration signal to M sub-signals whose frequency spectrum is as a PFS. The band-pass filter had a bandwidth similar to meshing frequency, where M is computed as the quotient of the division of the frequency range and the meshing frequency. The localized optimizing process of LADT improves the noise reducing capability in comparison with that of the ANR-GRS method; therefore, in this study it is used for denoising the vibration signal before the feature engineering process.

5.3.2 Wavelet-Based Vibration Imaging (WVI)

To obtain discriminant features from the preprocessed vibration signal, intrinsic information of the vibration signal should be utilized, such that it can provide enough information about MDTF types of defects. For this reason, a proper method that can highlight the key representative elements of MDTF-type defects in gearbox vibration signal is needed. Accordingly, the optimized output sub-bands from the LADT, which contains condensed defect-related useful information, are converted into two dimensional time–frequency representation images by employing the CWT method; these images are called WVIs. These WVIs, which carry enough fault-related information, are referred to as the enriched feature pool in this paper. The enriched feature pool of the WVIs can be utilized for identifying each defect type of MDTF states (i.e., PC, DT1, DT2, DT3, DT4, DT5, DT6) of the gearbox under variable speed conditions. The process of WVI formation can be explained in detail as follows.

To overcome the limitation of Fourier Transform in processing non-linear and non-stationary signals, and the limitation of STFT with fixed timing-window transforming observation, a wavelet approach has been developed. The wavelet transform uses a mother wavelet for decomposing a signal into the spatiotemporal domain. The mother wavelet can be adjusted by expanding or compressing during the transforming process [151]. We denote the wavelet function as $\varphi(t)$, with $\phi(\omega)$ as Fourier transform, to apply the wavelet approach in term of reversible transform, the admissibility condition must be satisfied.

$$C_\phi = \int_{-\infty}^{\infty} \frac{|\phi(\omega)|^2}{|\omega|}\, d\omega < \infty, \tag{5.4}$$

where C_ϕ is called admissibility constant. This (inequality 5.4) approximates that $\phi(\omega){=}0$, which can be presented as:

$$\int_{-\infty}^{\infty} \varphi(t)dt = 0, \tag{5.5}$$

and this requirement also makes clear that the mother function is a band-pass filter. The term 'wavelet' implies a small oscillation wave as the finite length of the window function, and "mother function" can be understood as prototype function such as Morlet wavelet or Daubechies wavelet, whose variants are the wavelet window functions. The actual wavelets are generated from a mother wavelet by the following equation:

$$\varphi_{s,\tau}(t) = |s|^{-\frac{1}{2}} \varphi\left(\frac{t-\tau}{s}\right), \tag{5.6}$$

where, τ is the translation parameter and s represents dilation in Equation (5.6). The translation parameter represents time in the wavelet domain. The dilation is the inversion of frequency. This scale of wavelet technique is analogous to the scale of map architecture. A large scale in mapping indicates the globalized scenery, and a smaller scale indicates more detail Similar principles can be applied to the wavelet approach; the high scale (i.e., $s \gg 1$, low frequency) is used for observing the global features of a signal because the wavelets are expanded for extracting the low-frequency components, such as the large time window of STFT. In contrast, the low scale (i.e., high frequencies, $s \ll 1$) is used for observing more details of a signal, called local features. Consider the vibration sub-band x(t) and the given wavelet family $\varphi_{s,\tau}(t)$, the continuous wavelet transform of $x(t) \in L^2(\mathbb{R})$ is calculated [152] by following inner products equation:

$$CWT_x^\varphi(s,\tau) = <x, \varphi_{s,\tau}> = |s|^{-\frac{1}{2}} \int_{-\infty}^{\infty} x(t)\varphi^*\left(\frac{t-\tau}{s}\right) dt. \tag{5.7}$$

Equation (5.7) represents the coefficients of *CWT*. *CWT* coefficients are the combination of translation series (time series) and scale (1/frequency) series, which can be utilized for constructing the vibration imaging feature spaces (scalograms). Through the use of the effective denoising technique from the previous process, the vibration image feature pools

are filled by condensed fault-related information that qualifies for the next identification step. The combination of the novel denoising technique and the CWT scalogram for the WVI are demonstrated in Figure 5.4 as the steps involved in the formation of WVI's.
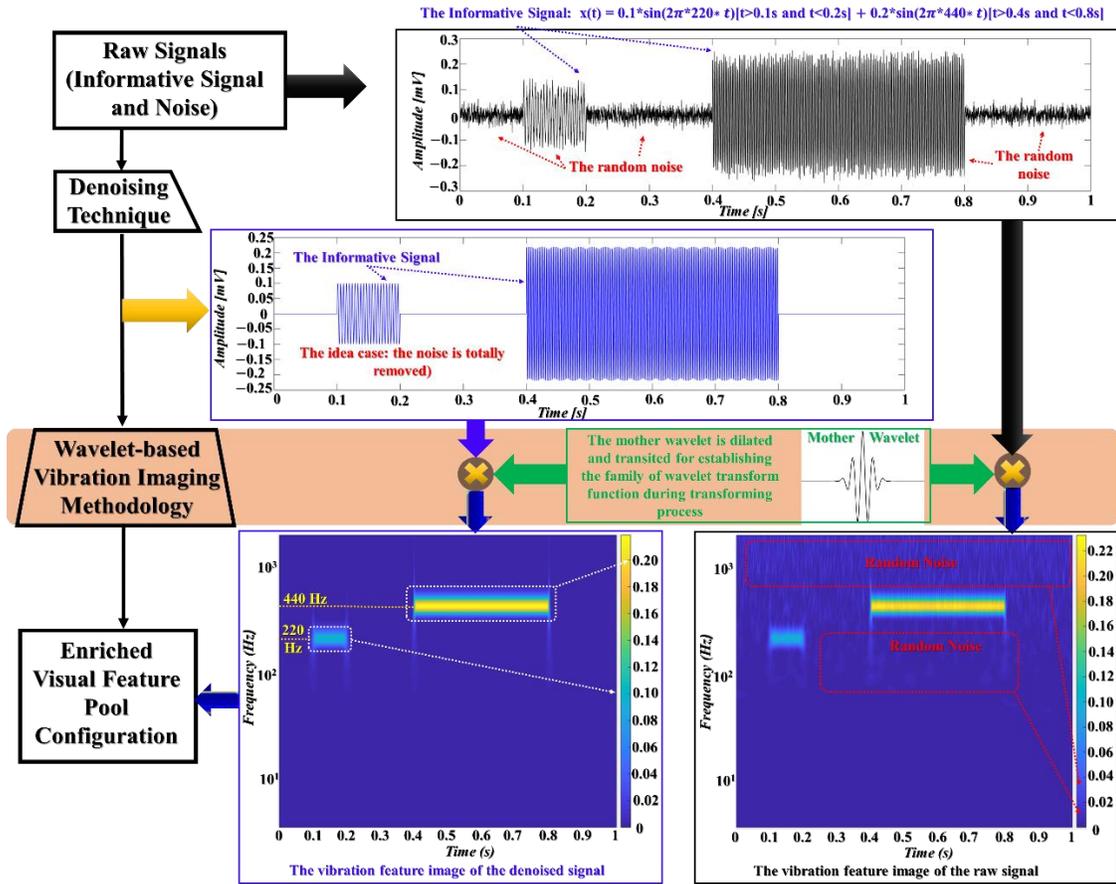
Figure 5.4. Steps involved in the construction of wavelet-based vibration imaging.

5.3.3 The Deep Convolutional Neural Network Architecture

DCNA comprises hidden layers (called convolutional layers), pooling layers, and fully connected layers [158][78]. The convolutional layer performs feature extraction from the input image data through a kernel-filter-based convolutional process; then, the pooling layer implements the down-sampling process. The pooling layer helps to reduce computational complexity and to recognize the learned extracted features. In addition, a variety of constraint-optimizing layers, such as rectified linear units (ReLU), dropout, and normalization, are integrated into the DCNA for classification improvement [163]. Afterward,

the fully connected layer uses weighted-base wiring to connect the output of the final convolutional or pooling layer for transferring information to the classification layer, which outputs the likelihood decision for classifying the fault types, normally using a SoftMax function [164]. Figure 5.5 demonstrates the general structure of the DCNA.

The convolutional layer (Cv) is responsible for the latent feature engineering processing. The Cv performs feature mapping through its layers for the extraction of representative attributes from input images that contain key information about gear states. To demonstrate the feature mapping process, we consider two consecutive layers as: j^{th} and $(j+1)^{\text{th}}$ convolutional layers, there are k filters (or kernels) with the size of $m \times n$, which are utilized for extracting feature from the output of jth layer. The output space of the jth layer, with the dimension of $m \times n$, is locally swept to convolve with each filter of $D \times R$ size using w training weights for adjustment. Then, each result, which corresponds to single kernel, is added in scale computation with bias b, and functionalized by activation functions of nodes in $(j+1)^{\text{th}}$ layer, these are normally non-linear functions, such as the rectified linear unit function (ReLU), used to perform non-linear feature mapping through layers. Assuming that the parameter used in the convolutional calculation is a unity, then feature space with a dimension of $(m - D + 1) \times (n - R + 1)$ is formed corresponding to each filter. In general, the i-th feature mapping space (fms) of the convolutional layer k can be formulated as follows:

$$fms_i^k = RL\left(\sum_{r \in A^{k-1}} fms_r^{k-1} \circledast w_i^k + b_i^k\right),\tag{5.8}$$

with RL is the ReLU function: $RL(x) = \max(0, x)$.

Where, w_i^k and b_i^k are the sets of weights and bias for i^{th} filter in layer k, $\circledast$ indicates the convolution operator, A^{k-1} denotes all feature mapping spaces in the $(k-1)^{\text{th}}$ layer. The feature spaces become more separable as it goes from lower convolutional layer to bottleneck layer network.

Typically, the pooling layer (Pm) is used next to each convolutional layer for the down-sampling process. It scans the whole range of a feature mapping space sequentially, and then applies the pooling operation on a defined pooling region by a non-overlapping searching method. The pooling operation that is most commonly used is the mean average, or maximum value in the defined pooling area [158].

Usually, many incorporated pairs of convolutional and pooling layers are employed in DCNA. After the final convolutional layer or pooling layer, several fully connected layers (Fc) are used to expand deep representation feature mapping spaces, as well as the concatenation of feature mapping spaces into a feature vector. Finally, the represented feature vectors are provided as an input to non-linear nodes for classifying the features into their corresponding categories (the fault states of a gearbox). The SoftMax function is typically used as the final activation function in the classification layer for classifying the input data into their corresponding categories.
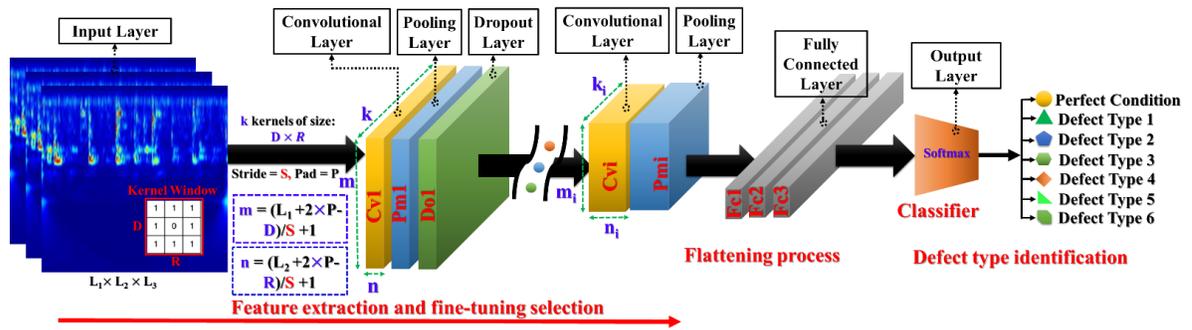
Figure 5.5. Description of typical DCNA

The learning process of the DCNA is based on the optimization of the loss function of the reconstruction error. The loss function is the function of the training error, which is the difference between predicted output ($\hat{y}_q$) and actual output (y_q), it can be presented as follows:

$$e(n) = \frac{1}{2}\sum_{q=1}^{K}(y_q^n - \hat{y}_q^n)^2. \tag{5.9}$$

Here, K signifies the number of neurons, n is the order of repetitive steps. The major purpose of the training process in building of the DCNA to fine-tune its parameters converging to reduce $e(n)$ through a back-propagation process based on the stochastic gradient descent method [165].

5.4 The Accurate and Stable MDTF Fault Identification Framework and Its Experimental Evaluation

The key aim of this study was to identify defect types of MDTF gearbox systems under variable speed conditions. As mentioned in section 5.1, it has been observed that the existing models might not be able to differentiate those fault types due to the similar behavior of different degrees of tooth fault reflected in the vibration spectrum. To address this issue, in this paper, a new gearbox fault diagnosis scheme has been proposed. Figure 5.6 provides a block diagram of the proposed framework. From Figure 5.6 it can be seen that the proposed method consists of four main steps: (1) sensors and data acquisition (DAQ), (2) LADT, (3) WVI, and (4) DCNA. The preliminary section covered the main steps of the proposed method. This section will provide the experimental validation of the proposed method.
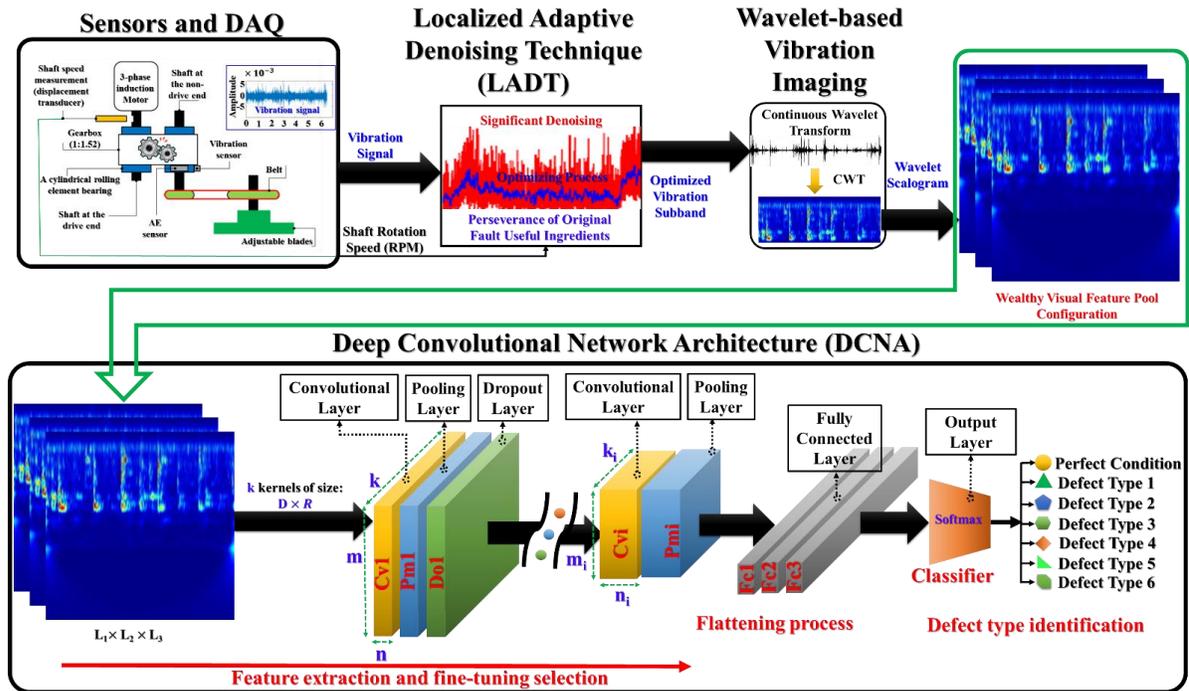
Figure 5.6. Block diagram of the proposed accurate and stable MDTF gear fault identification framework.

5.4.1 The Gearbox Testbed and Data Acquisition

A gearbox testbed, self-developed at the Ulsan Industrial Artificial Intelligence laboratory, for acquiring vibration data is shown in Figure 5.7. The testbed can be explained as follows: an AC motor is directly connected to the pinion wheel through the drive shaft (DS), whereas the gear wheel is fixed with a non-drive shaft (NDS) and the adjustable blades (the load). The pinion wheel with 25 teeth, whose length is 9 mm, and the gear wheel (38 teeth) are engaged with each other and housed in the gearbox, creating a gear reduction ratio of 1:1.52. The rotation movement (torque) of the load is provided by the motion of the AC motor through the gearbox. Therefore, the rotational speed of the pinion wheel is equal to the rotational speed of AC motors, and the gear frequency is calculated by the pinion frequency and the gear ratio. The vibration sensor (the accelerometer) is placed at the end of the NDS, 72.5 mm from the gear wheel. The rotational speed of the DS (a pinion frequency) is measured by the displacement transducer, which is mounted for tracking the hole in DS once per revolution. The data acquisition system, which is the PCI-2 data acquisition board, is connected to the accelerometer (622B01) to measure and digitize vibration signals, and to store digital vibration samples. The specifications of the accelerometer, speed sensor, and data acquisition system are given in Table 5.1.
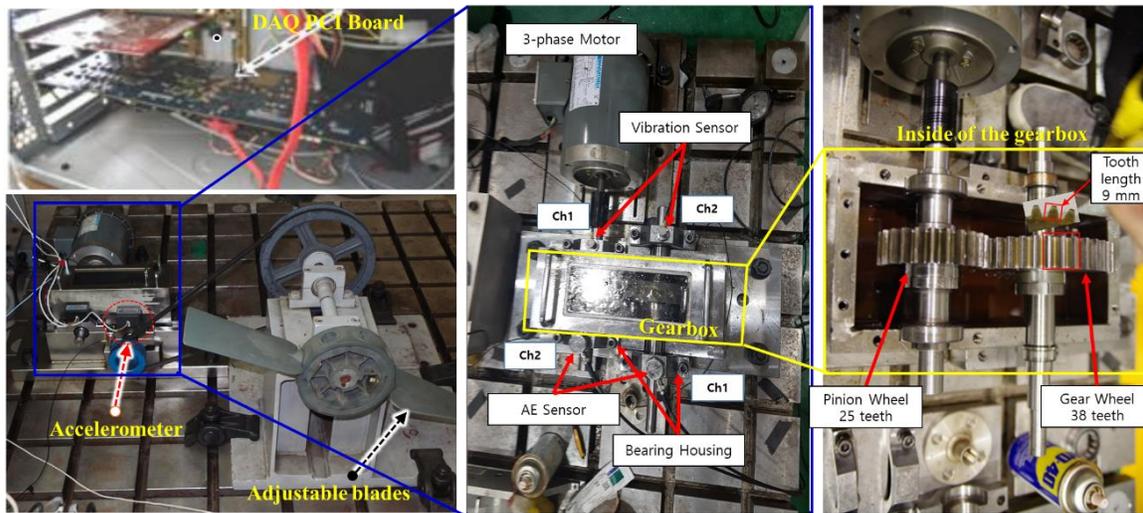
Figure 5.7. Gearbox experimental testbed

Table 5.1 Specification of the sensors and data acquisition system.

Devices	Specification
Vibration sensor (Accelerometer 622B01)	Sensitivity (V/g): 10.2 mV/(m/s^2)
	Operational frequency range: 0.42 to 10 kHz
	Resonant frequency: 30 kHz
	Measurement range: ±490 m/s^2
4- Channel DAQ PCI Board	18-bit 40MHz AD conversion, a sampling frequency of 65.536 kHz is used for each of two channels simultaneously
Displacement transducer	Distance from the head of a transducer to a hole: 1.0 mm
	Diameter of a hole: 12.80 mm
	Sensitivity: 0 to -3dB
	Frequency response: 0–10 kHz

The MDTF gearbox was created by cutting one tooth, mounted on the gear wheel, to different degrees. Figure 5.8 shows the degrees of cut teeth and the vibration signals obtained under each condition for all observed defect types in this study, including a normal or perfect condition gear (PC), 6.6% degree of tooth defect (DT1), 10% degree of tooth defect (DT2), 20% degree of tooth defect (DT3), 30% degree of tooth defect (DT4), 40% degree of tooth defect (DT5), and 50% degree of tooth defect (DT6). These multiple degree tooth faults were seeded for simulation of the same behavior of the gear defects caused by long-term operation of a gearbox system (e.g., tooth spalling, tooth cracking, worn tooth, etc.). The vibration characteristic for fault states of a gearbox was analyzed in detail in Section 5.2.
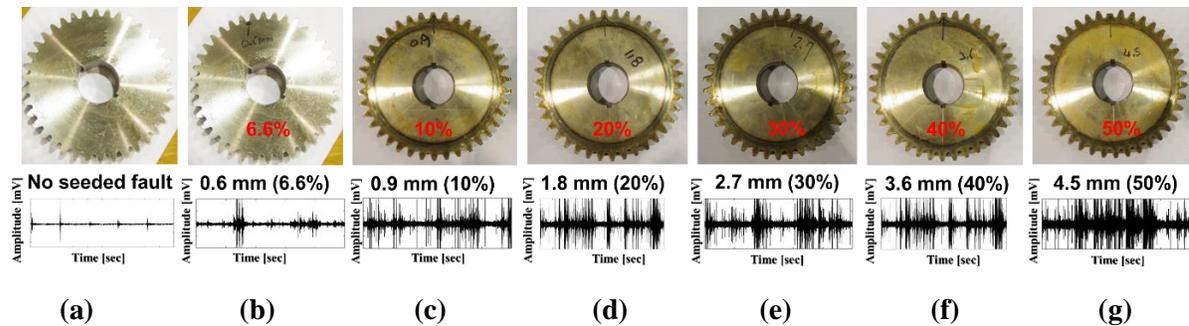
Figure 5.8. The observed defect types of the multi-degree tooth faults on the gear wheel and examples of vibration signals at 600 RPM: (a) PC, (b) DT1, (c) DT2, (d) DT3, (e) DT4, (f) DT5, and (g) DT6, respectively.

Table 5.2 demonstrates the configuration of the dataset used in this paper. The data acquisition system converts the analog vibration signal to a digital vibration signal with a sampling frequency of 65,536 Hz. Each sample is one second long, termed a one-sec sample. A total of 200 samples were collected under each defect condition with variable rotating speed (four shaft rotational speeds are evaluated in this study). Therefore, there are 800 samples for each defect condition, and a total of 5600 samples are extracted from this testbed.

Table 5.2. The configuration of the MDTF gearbox dataset.

Gearbox Defect Type	Description	Number of one-second samples for specific rotation speed (RPM)				Sampling frequency (Hz)
		300	**600**	**900**	**1200**	
Perfect Condition (PC)	Normal or Perfect gearbox	200	200	200	200	65,536
Defect Type 1 (DT1)	6.6% degree of tooth defect (0.6 mm/9 mm)	200	200	200	200	65,536
Defect Type 2 (DT2)	10% degree of tooth defect (0.9 mm/9 mm)	200	200	200	200	65,536
Defect Type 3 (DT3)	20% degree of tooth defect (1.8 mm/9 mm)	200	200	200	200	65,536
Defect Type 4 (DT4)	30% degree of tooth defect (2.7 mm/9 mm)	200	200	200	200	65,536
Defect Type 5 (DT5)	40% degree of tooth defect (3.6 mm/9 mm)	200	200	200	200	65,536
Defect Type 6 (DT6)	50% degree of tooth defect (4.5 mm/9 mm)	200	200	200	200	65,536

5.4.2 LADT Performance for Effective Noise Removal of Vibration Signals of a MDTF Gearbox under Variable Speed Conditions

The raw vibration signals were digitized at a high sampling frequency of 65,536 Hz in order to gather rich discrete vibration samples, and to capture the extent of feasible defect-related components in each one-sec vibration signal. The vibration data collected from the gearbox contain fault-related information and interference noise. By sampling the vibration signal at a frequency of 65,536 Hz, the frequency spectrum of a discrete vibration sample is from 0 Hz to 32,768 Hz (according to the Nyquist–Shannon sampling theorem). However, the accelerometer is capable of sensing the vibration oscillations in the frequency range of 0.42–10,000 Hz (Table 5.1). Thus, the fault-related information is in the frequency range of 0.42–

10,000 Hz. Therefore, rather than providing the raw vibration signal to LADT, the vibration signal is pre-processed by performing down-sampling using a low-pass filter to avoid aliasing [126]. After performing down-sampling, the vibration sub-bands are obtained, which have the time length of one second, the sampling frequency of 21,845 Hz (65,536/3), and frequency range from 0–10,922 Hz.

The vibration sub-bands are provided as an input to LADT for reducing the noise and enhancing the useful fault-related information, which represents multi-degree tooth fault behaviors. LADT applies ANR-GRS to each PFS. Through localized adaptive optimization, the new denoising methodology tries to find the localized optimal parameter set of the noise-simulated reference signal, which is appropriate to each specific PFS. The outputs of LADT are the optimized vibration sub-bands, which maintain the original defect-related frequency tones (meshing frequency harmonics and sideband frequencies) and reduced background noise. Those defect representative ingredients are key factors for identifying the defect types of an MDTF gearbox under the condition of the variable speed. Those fault types proceed as analogous behaviors reflecting the vibration characteristic, the differences between them might be degrees of amplitudes of informative tones, their proportions, or occurrence events in the tiny range of separation. Thus, the image-based enriching feature pool configuration methods are needed to sort them out for condensation.

5.4.3 Wealthy Feature Pool Configuration Based on VWI

In this step, a continuous wavelet transform is applied to the noise-free optimized sub-bands obtained from the LADT. The wavelet-based transforming method is used to convert time-domain optimized sub-bands to the scalograms for the enriched visualized features pool. The CWT method in this paper employs the Morlet wavelet, which is the most effective technique in the fault diagnosis approach of the rotation machines [166], as a mother wavelet function. The signal is decomposed up to 16 octaves. Based on experiments, the optimal value of voices per octave parameter was chosen as 16. The wavelet coefficients for each input sub-band, which are derived from applying the wavelet family functions in Formula (5.5), are used to obtain a scalogram, which is an energy distribution map of the input sub-band on a time–frequency scale. Those scalogram images of the vibration sub-bands are reshaped by the size

of 224×224×3 for compatibility with the input layer of DCNA in the next classification step and packed to configure the enriched visual image feature pool.

5.4.4 DCNA Construction

In this study, the contents of the enriched visualized feature pool, which are called WVIs are obtained from the CWT of low-noise optimized vibration sub-bands, are provided as an input to DCNA. The WVI contains fault-related information in the form of edges, lines, curves, spots, or pixels with various intensities (which are represented by the R, G, and B channels of the RGB image). The DCNA is used primarily to recognize images. Figure 5.9 demonstrates the architecture of the proposed DCNA used in this study. The proposed DCNA has fifteen layers, including five convolutional layers (Cv), three pooling layers (Pm), two drop-out layers (Do), three fully connected layers (Fc), one input layer, and one terminal output layer (Os). The DCNA makes a start with an input layer of size $224 \times 224 \times 3$, according to the size of RGB images (224×224 indicates the values of length and width, and 3 denotes three R, G, B channels of the input image). Next, the features are extracted from fault-related images by the first convolutional layer with 96 kernels of size $11 \times 11 \times 3$ and the stride of 4. The results of the first convolutional calculation are feature spaces of size 54×54×96. After the first convolution layer (Cv1), the max-pooling layer (Pm1) is applied for down-sampling. Moreover, the drop-out layer (Do1) is located in series to extenuate the over-fitting issue [163]. The second convolutional layer has 256 filters of size $5 \times 5 \times 48$, and it is followed by pooling and dropout processing layers. The Cv3 and Cv4 layers consist of 384 filters with a size of $3 \times 3 \times 256$ and 384 kernels with a size of $3 \times 3 \times 384$, respectively. Next, Cv5 is down-sampled by the third max-pooling layer (Pm3), composed of 256 of $3 \times 3 \times 384$ kernels. All of the max-pooling layers employ 3×3 filters with a stride of 2. The output of the third max-pooling is used as an input to the fully connected layers (Fc1, Fc2, Fc3). Fc1 tries to implement a flattening process to convert all feature matrices ($6 \times 6 \times 256$) from the output of layer Pm3 to the feature vectors ($1 \times 1 \times 4096$) through its operation as a weighted sum with bias terms. These output feature vectors then are passed through the activation function ReLU and input to the next layer (Fc2). The second fully connected layer, which is the penultimate layer, includes 1000 neurons and functions, similarly to Fc1, to output feature vectors of size $1 \times 1 \times 1000$. The last flattened layer, Fc3, including 7 neurons, which are the

SoftMax activation functions, is the classification layer. It operates at a terminal spot of the DCNA for estimation of the probabilities of the categories.
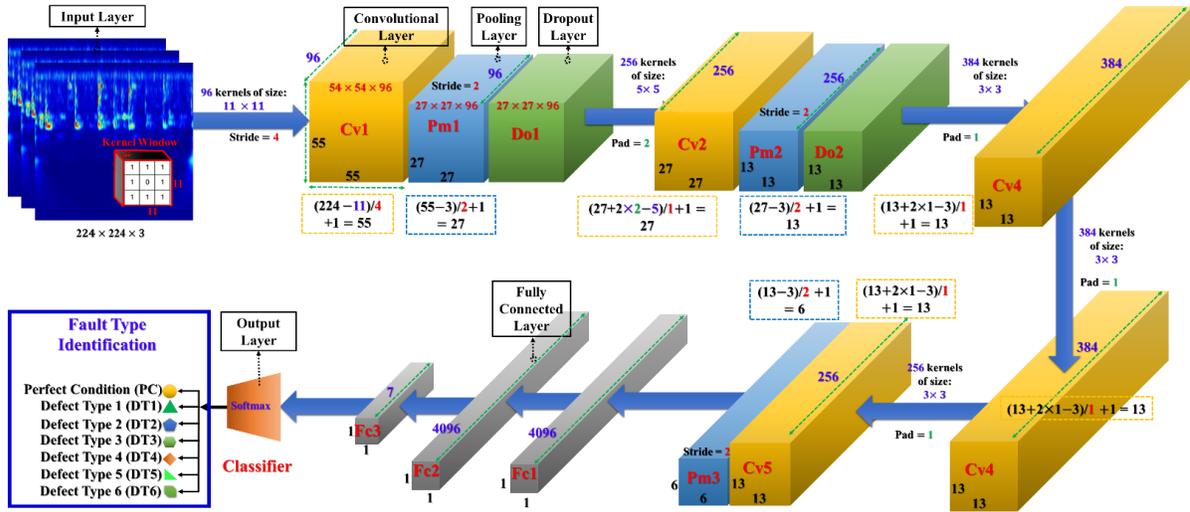
Figure 5.9. The applied DCNA model for implementing the fault type identification in this study.

In this paper, the fifteen-layer DCNA has been conducted based on the original AlexNet architecture [167], with some modifications for this specific application. The AlexNet model has already achieved better feasibility than other models for recognizing images. This model has implemented training for 1.2 million high-resolution pictures of ImageNet for classification of up to 1000 differential species targets in the contest of LSVRC-2010 by training of 650 thousand neurons and 60 million parameters, with many optimizing processes in the network architecture. In our research, we have replaced two normalization layers with two drop-out layers in order to improve the capability of over-fitting avoidance [163][168]. Moreover, the last fully connected layers (Fc3), which include 1000 neurons from the original AlexNet, are replaced by the same fully connected layers with a reduced number of neurons (7), for suitable application in our research with seven classifying categories. The detailed description of the proposed DCNA is shown in Table 5.3.

Table 5.3. The structural elements of the proposed DCNA.

Layers	Operating Parameters	Number of Kernels	Kernel Size	Stride	Padding
Input layer	224×224×3				
1st Convolutional (Cv1)	55×55×96	96	11×11	4	0

1st Max Pooling (Pm1)	27×27×96	96	3×3	2	0
1st Dropout (Do1)	27×27×96				
2nd Convolutional (Cv2)	27×27×256	256	5×5	1	2
2nd Max Pooling (Pm2)	13×13×256	256	3×3	2	0
2nd Dropout (Do2)	13×13×256				
3rd Convolutional (Cv3)	13×13×384	384	3×3	1	1
4th Convolutional (Cv4)	13×13×384	384	3×3	1	1
5th Convolutional (Cv5)	13×13×256	256	3×3	1	1
3rd Max Pooling (Pm3)	6×6×256	256	3×3	2	
1st Fully Connected (Fc1)	1×1×4096				
2nd Fully Connected (Fc2)	1×1×4096				
3rd Fully Connected (Fc3)	1×1×7				
Output	SoftMax Nodes				

5.4.5 The Experimental Classification for an MDTF Gearbox under Variable Speed Conditions

The DCNA performs a fault-classifying process based on the input WVI imaging data for the MDTF gearbox under varying speed conditions. To verify the performance of the proposed DCNA for identifying seven MDTF fault types under varying speed conditions, we conducted an experimental setup of two scenarios, as shown in Table 5.4. In Scenario 1, all vibration data for four speeds were observed for classification. While in Scenario 2, four experiments were performed based on varying speed-related data. The configuration of the testing and training datasets for both of the scenarios is given in Table 5.4.

Table 5.4. Description of the dataset for training and testing with RPM in the experiment setup.

Scenarios	The experiment	Number of samples	The RPM of data samples
Scenario 1	Experiment 0	Training samples: 3360	60% of All four speeds dataset
		Testing samples: 2240	40% of All four speeds dataset
Scenario 2	Experiment 1	Training samples: 2800	The shaft speeds: 300 RPM, 600 RPM
		Testing samples: 1400	The shaft speed: 900 RPM
	Experiment 2	Training samples: 2800	The shaft speeds: 600 RPM, 900 RPM
		Testing samples: 1400	The shaft speed: 1200 RPM
	Experiment 3	Training samples: 2800	The shaft speeds: 900 RPM, 1200 RPM
		Testing samples: 1400	The shaft speed: 300 RPM
	Experiment 4	Training sample: 2800	The shaft speeds: 1200 RPM, 300 RPM
		Testing samples: 1400	The shaft speed: 600 RPM

For each speed (a total of four speeds: 300 RPM, 600 RPM, 900 RPM, and 1200 RPM), there were a total of 1400 one-second samples for all gear fault types (there were seven defect types

or seven categories, PC, DT1, DT2, DT3, DT4, DT5, DT6, and each of them was acquired by sampling for one second, repeated 200 times, to achieve 200 one-second samples). All these samples were first preprocessed using LADT. Next, the output optimized sub-bands obtained from LADT were converted by the CWT method to attain the enriched feature scalogram images. That speed-related image subset was used as input data for the DCNA. For each experiment, we used two speed-related datasets (2800 samples) to train the proposed DCNA several times with multitudinous epochs, targeted to optimize the network parameters based on minimizing orientation of the loss function (Equation (5.9)), and the dataset of another speed (1400 samples) was used as the testing dataset of the constructed model. These processes were circularly acted based on four speed-related datasets to conduct all four experiments.

5.5 Results and discussion

This section principally validates the proposed fault identification framework constructed in section 5.4 for an MDTF gearbox under inconsistent rotational speeds based on the data collected from a real-world testing platform. The effectiveness of this model is entirely evaluated based on the following operations: LADT, visual enriching feature configuration (WVI's), and fault identification based on DCNA.

5.5.1 Experimental Verification of the Effective Performance of LADT and Wealthy Feature Pool Configuration Created by WVI

As explained in the introduction section, the real-world gearbox vibration signals originally contain informative components and random background noise. The disturbance noises appear randomly, and they can affect the informative components. Thus, in the raw form of the vibration signal, it is very difficult to separate the original informative components from the background noise. Furthermore, the operation behaviors of MDTF gear faults reflected in the vibration signal are too similar. In other words, to discriminate these kinds of faults, the use of enhanced techniques is required. The LADT approach is the key technique of this study for effective noise cancellation and for separating the original fault-related components from the high noise vibration signals. Before being fed to LADT, the raw vibration signals, gathered

from the experimental gearbox testbed, were processed by down-sampling and low-pass filtering to attain vibration signals with the frequency range of 0–10,922 Hz, according to the real frequency working range of the acceleration sensor for removing the redundancy fractions. These output signals are named raw-filtered vibration signals. LADT tries to divide each raw-filtered vibration signal into many sub-signals, so that their frequency spectrums are principal frequency segments, by applying the series of the non-overlapped band-pass filters along the frequency spectrum of the vibrations signal (0–10,922 Hz). Next, the ANR-GRS technique [126] is applied on each principle frequency segment to achieve a locally optimized sub-band from each input sub-signal based on the localized optimal parameters. The final optimized output of the LADT module is a summation of all locally optimized sub-bands corresponding to each input vibration signal.

The visual analysis of frequency spectrums of three vibration signals (a raw-filtered vibration signal, the output signals of the ANR-GRS module, and optimized sub-bands from the LADT module) are illustrated in Figure 5.10. As shown in Figure 5.10, the superiority of the localized adaptive process of the LADT module for denoising is proven. Here, the noise disturbance areas, which were circled by red dotted lines in the spectrum of a raw-filtered vibration signal that inputs to ANR-GRS and LADT modules, were mostly removed in the output of the ANR-GRS module and LADT module (the spaces with red narrows in the output spectra of ANR-GRS and LADT modules). However, the output signal of LADT indicated outstanding efficiency in reducing noise relative to the ANR-GRS module; the noise areas of the second and fifth principal frequency segments (the segment contains the second and fifth harmonics of the meshing frequency) in the output sub-band of the LADT module were much lower than those in the output signal of ANR-GRS. This verifies the effectiveness of the localized adaptive optimization process of the LADT scheme. In addition, the fault-related components, which were marked by blue-dotted circles in the input and output of LADT, were exactly the same. In other words, the LADT approach reduces noise in the largest amount possible by obeying the principal rule of a condition-monitoring fault diagnosis system to preserve the original fault-informative elements, such as sideband frequency tones and meshing frequency harmonics inside of the raw vibration signals.
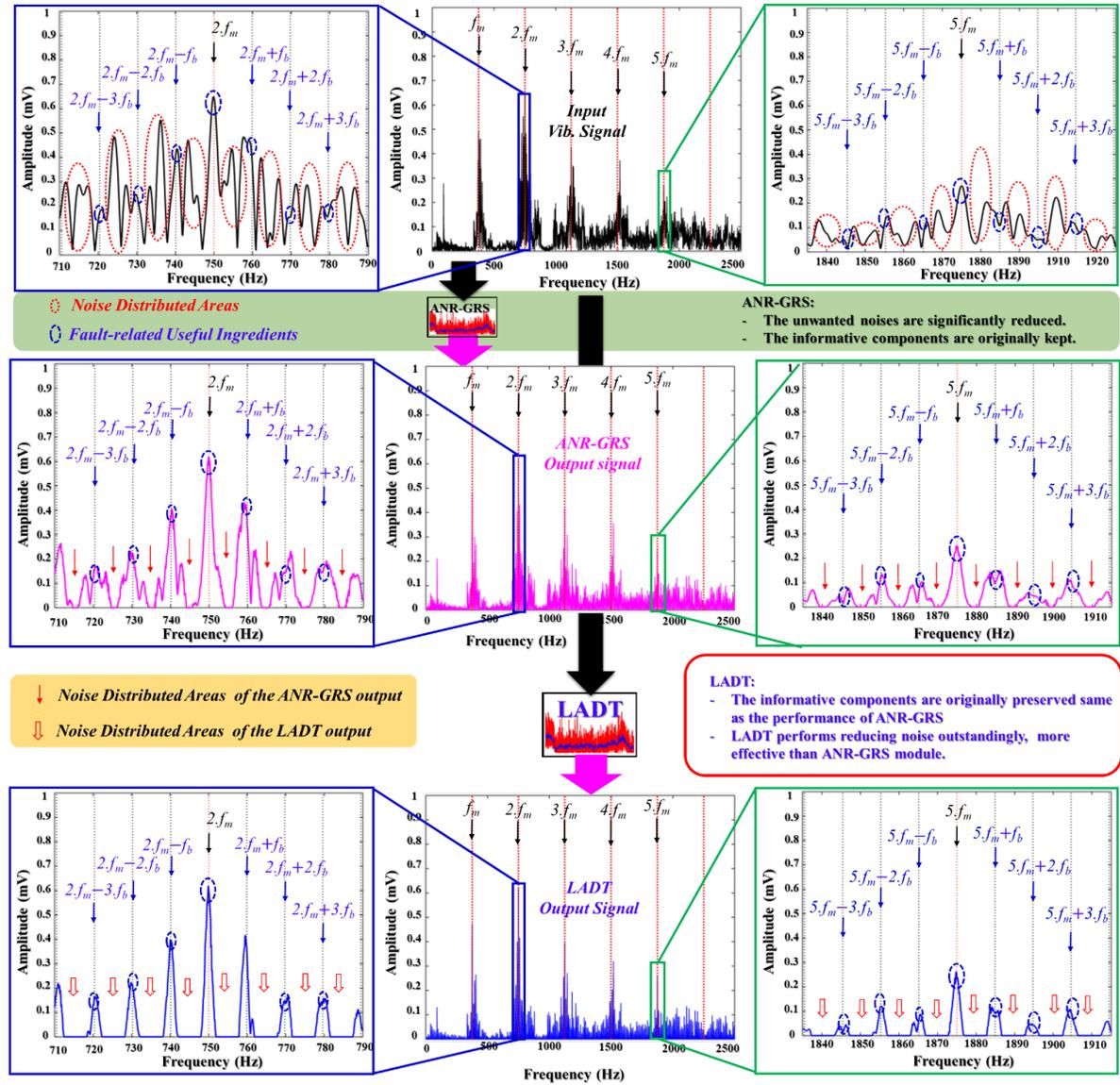
Figure 5.10. The frequency spectrum analysis of the input and output signal of LADT in comparison with the performance of ANR-GRS for an example vibration signal of DT3 at 900 RPM.
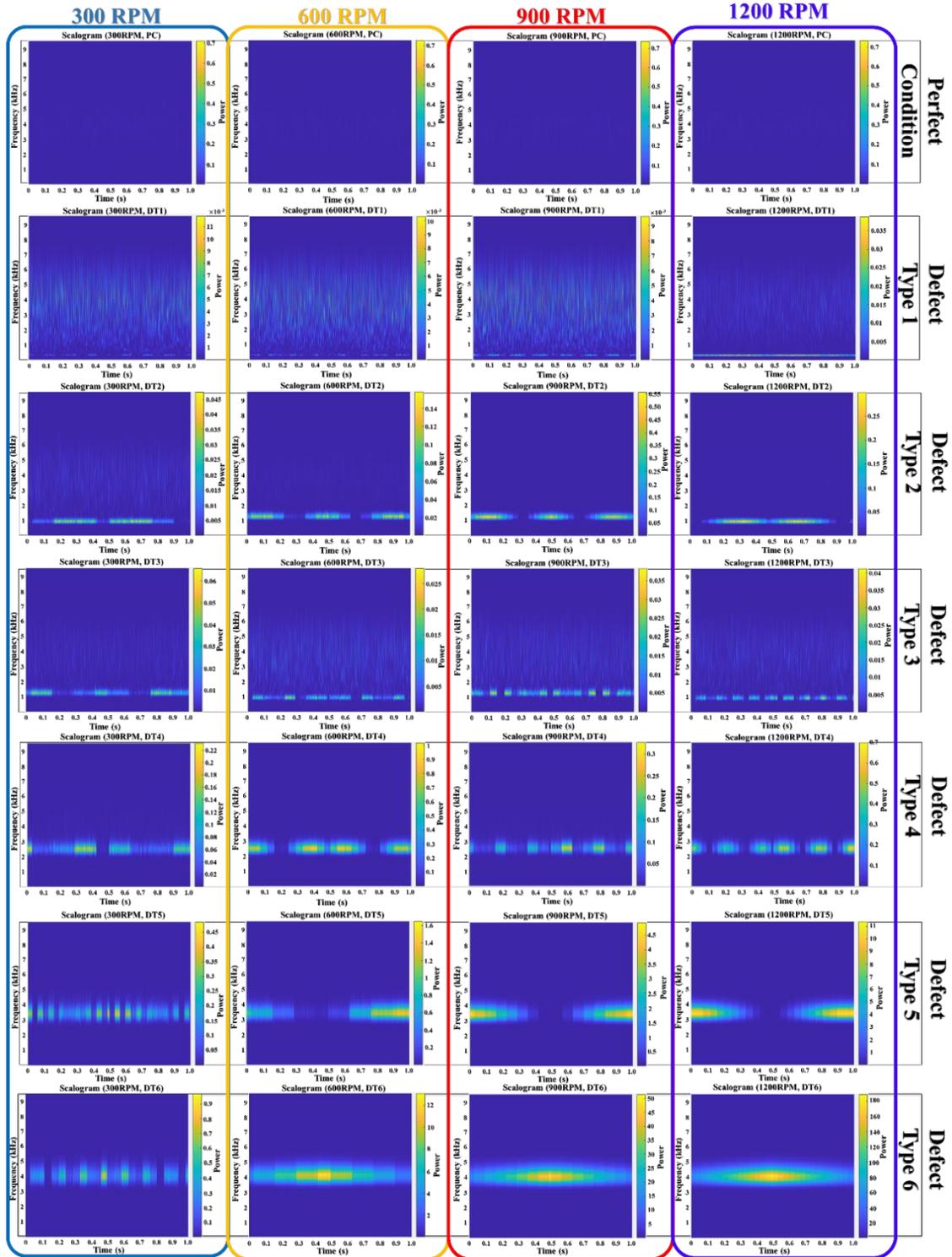
Figure 5.11. Frequency spectrum analysis of the vibration subband (for fault state D2 at 900 RPM) in the comparison between an input and output subband of the ANC module.

The output optimized sub-bands from LADT were then converted to visualized feature spaces,

for better expression of defect-related components induced by vibration characteristic of MDTF defect types in the time–frequency domain, using the proposed WVI method. Similarly, to the example signal in Figure 5.4 (section 5.3.2), the wavelet-based vibration images carried the defect-correlated factorials and exposed the attributes through color images. Figure 5.11 demonstrates the scalograms of the seven defect types of a gearbox under four rotational speeds. Through visualization, the scalograms of the same defect type under four rotational speeds showed the proximate parallel zones with the different energy levels. In addition, the energy of the useful components (pixel illuminations) has grown according to the uptrend of rotational speeds. Those discriminate notifications were quantized in the massive process of feature extraction and optimization achieved from DCNA performance.

5.5.2 DCNA-Based Identification Performance Analysis

By applying the LADT method, the noise components of the vibration signals were mostly removed. The wealthy feature pool configuration based on CWT, then, translated the output of LADT as insignificant-noise vibration sub-bands to the scalogram images. These scalogram images carried enough information for fault discrimination. The wavelet-based vibration image datasets were used as input datasets for DCNA for the classification task. First, the proposed network tried to perform Scenario 1 to discover the effect of the quantity of input data on the time consumption and classification accuracy. The dataset, which contained all four speeds and seven categories, was randomly split into the training set and validation set. Each input sample was a colorized image with dimensions of $224 \times 224 \times 3$, which met the demand of the input layer size of the proposed DCNA. From the numerous proportions of the training set, the computational consumptions and accuracies are listed in Table 5.5. It shows that when 50% to 60% of total samples were used for training, the best performances were obtained (by high accuracies in the acceptable time consumption) in the observed quantities. Thus, a ratio of 60% was used in this study.

Table 5. 5. The classification accuracy and time consumption for various size of the training set.

Training size (Percentages of 5600 samples)	Number of Epochs	Time consumption (second)	Overall Classification Result (%)
1680 Samples (30%)	160	105.101	89.51
2240 Samples (40%)	200	138.276	94.63

2800 Samples (50%)	210	147.846	99.79
3360 Samples (60%)	250	165.569	100
3920 Samples (70%)	300	375.497	100
4480 Samples (80%)	360	458.990	100
5040 Samples (90%)	410	546.832	100

In scenario 2, four experiments (in Table 5.4) were executed in this study to analyze the accuracy and reliability of the proposed framework for an MDTF gearbox under differential speed conditions. In each of the four experiments, the training dataset was composed of two different speed samples (2800 samples), and the data samples of the validation set contained samples collected at speeds that differed from that of the training dataset (1400 samples). Following Scenario 2, the speed-varying datasets were alternately used for the training and testing process over a total of four observed rotational speeds in this paper. The learned features of the activation processes in different layers of the applied network model can be seen in Figure 5.12. From the input RGB image of the defect type 3 with a speed of 600 RPM (Figure 5.12a). through the beginning steps of the high-dimensional feature extraction process, performed by 96 kernel filters (Figure 5.12b) of the first convolutional layer (Cv1), the feature images of the Cv1 of one channel are shown in Figure 5.12c. With the help of this process, the one time-frequency domain vibration image is mapped to 96 feature images for observing the defect-related elements in high-dimensional feature spaces. Next, the several mapping values in feature images are reduced by the max-pooling layer (Pm1) as shown in Figure 5.12d. Thus, the feature image in Figure 5.12d is inclined viewed-dubiously and softer than Figure 5.12c. From Figure 5.12e to Figure 5.12h, the complex learned feature images from Cv2 to Cv5 of an example channel are demonstrated the impacts of the kernels of those layers. After flowing through Cv and Pm layers of the applied DCNA network, the learned feature maps were flattened as feature vectors. Those feature vectors, which were outputs of the final fully connected layer (Fc3), were then used as input of a SoftMax layer or output layer for clustering.

The t-SNE (t-stochastic neighbor embedding) approach is popular in deep networks for exploring the feature spaces. Figure 5.13 depicts the three-dimensional distribution of the output feature vectors from the Fc3 layer according to seven defect categories through four experiments. As shown in Figure 5.13, the samples of the same defect type were close to each other, separate from the samples of another defect type. The clear discrimination between

defect types verifies the high accuracy and stable capability of the proposed framework through the condition of the inconsistent speed. Based on this, the classification process can identify the defect types of an MDTF gearbox more easily.
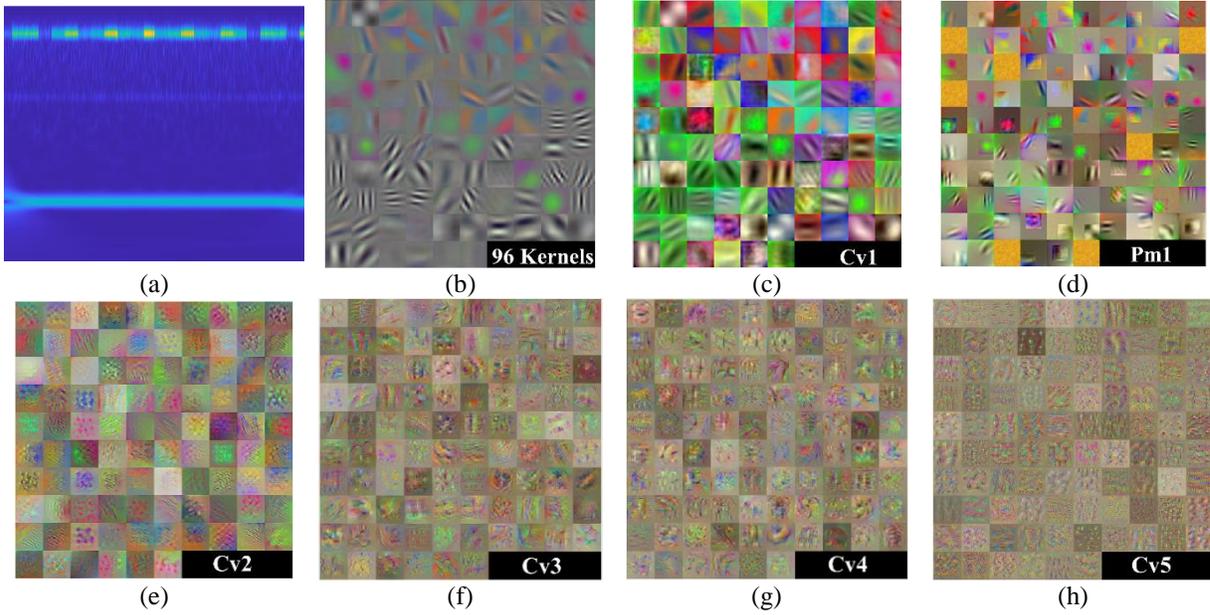
Figure 5.12. The flowing learned feature images through layers of the proposed DCNA for one example channel, here, (**a**) RGB input image, (**b**) the 96 kernels of size 11×11 (**c**), the feature images of the Cv1, (**d**) the feature images of the Pm1, (**e**) the feature images of the Cv2, (**f**) the feature images of the Cv3, (**g**) the feature images of the Cv4, the feature images of the Cv5 (**h**).
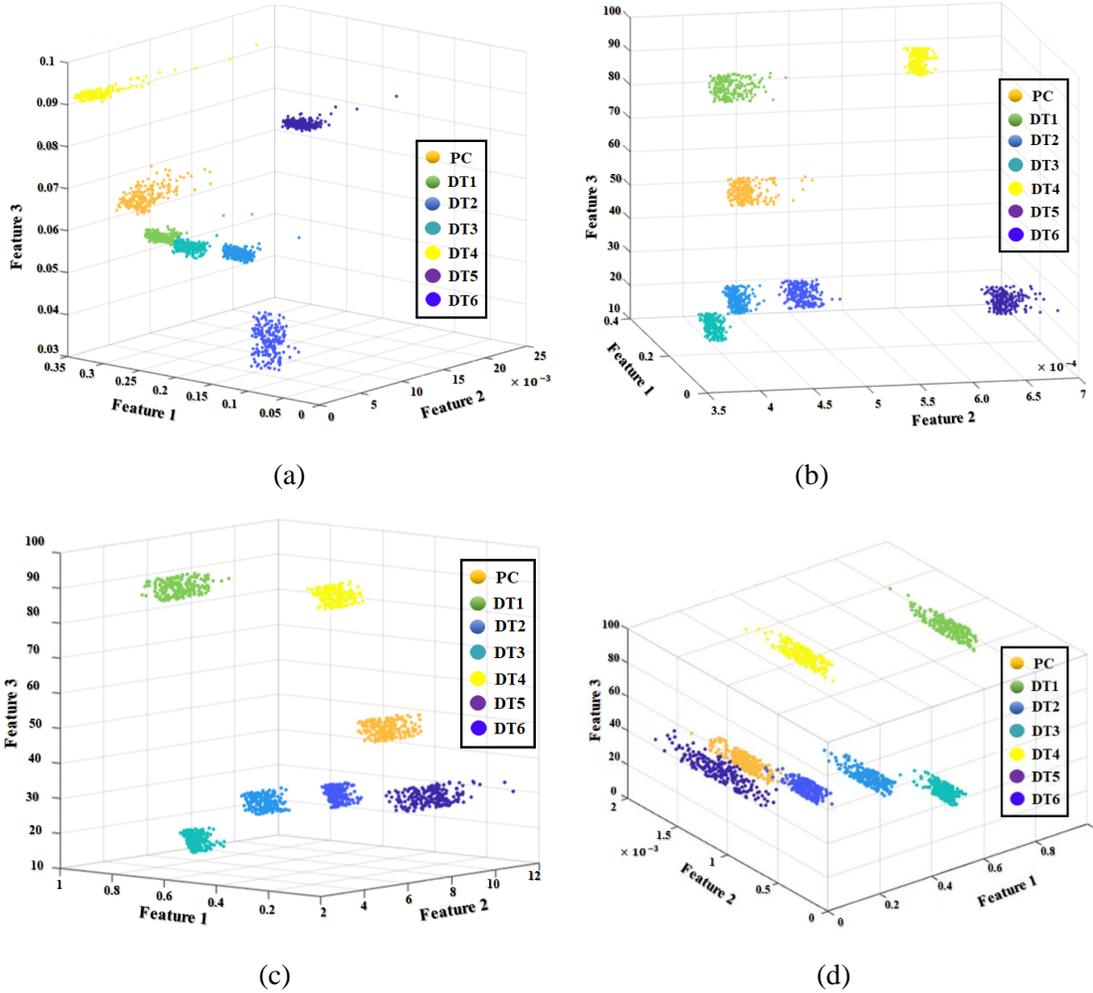
Figure 5.13. Three-dimensional clustering spaces of the four experiments: (a) experiment 1, (b) experiment 2, (c) experiment 3, (d) experiment 4.

Moreover, the confusion matrix, which is shown in Figure 5.14, provided perfect performance (100% accuracy) of fault identification for seven defect types of the experimental MDTF
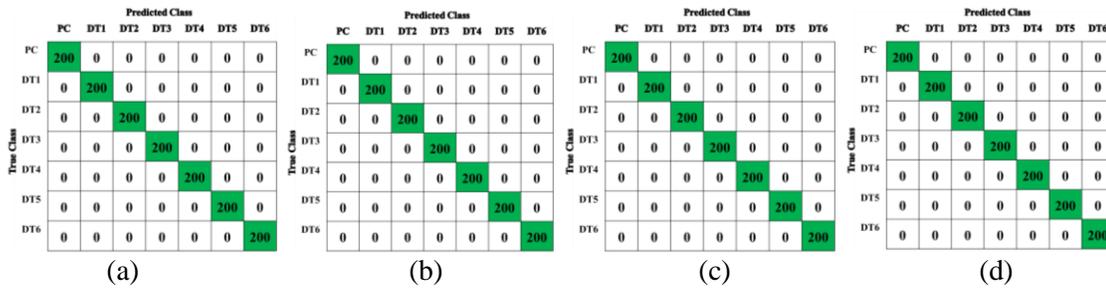
Figure 5.14. The confusion matrices of the experimental scenario 2: (a) Experiment 1, (b) Experiment 2, (c) Experiment 3, (d) Experiment 4.

gearbox under variable speed conditions through the four experiments in Scenario 2.

For robustness analysis of the proposed methodology, a comparison was made between the proposed method and existing state-of-the-art methods such as ANR-GRS + SFE + GA + KNN (Fw1), LADT + GA + KNN (Fw2), LADT + FSE + SVM (Fw3), ANR-GRS + CWT + DCNA (Fw4), LADT + STHT + DCNA (Fw5). Those are explained in detail as follows:

(1) ANR-GRS + SFE + GA + KNN (Fw1): This framework used the denoising method as an adaptive noise-reducer-based Gaussian reference for optimizing vibration signals. Next, the handcraft feature extraction technique was used to extract the statistical features in the time and frequency domain (SFE: statistical feature extraction). The achieved feature pool, then, was processed by a feature-selection-method-based genetic algorithm (GA) to fetch the most discriminate features in preparation for input into the learning model as k-nearest neighbors (KNN). KNN performed fault classification based on the selected features (reduced dimensionality) to identify the gear defect types for validating the accuracy of the constructed model (Fw1). The details of Fw1 can be found in [169].

(2) LADT+ GA + KNN (Fw2): To validate the improved denoising technique, the LADT module was used instead of the ANR-GRS module in the Fw1 to construct the Fw2.

(3) LADT + FSE + SVM (Fw3): This observed framework was created to explore the noise reduction proficiency of LADT, incorporating the high-dimensional feature pool, which can be well-classified by a support vector machine (SVM). The proposed denoising approach (LADT) in this study was applied to optimize vibration signals. The FSE step tried to configure the feature pool. Then, an SVM was utilized to process fault diagnosis by using the extracted features to input learning data [126].

(4) ANR-GRS + CWT + DCNA (Fw4): By implementing this framework, the effectiveness of the LADT module was straightforwardly compared to the initial adaptive noise technique (ANR-GRS). In this situation, we only replaced the LADT module with ANR-GRS.

(5) LADT + STHT + DCNA (Fw5): This framework was implemented by using short-time Fourier transform (STHT) to extract the visualized image features as spectrogram

images. It was used for comparison with the proposed scheme in the process of enriching feature extraction.

Those methodologies were selected to evaluate the performance of the proposed method in terms of the improvement of LADT for denoising in comparison with the initial method (ANR-GRS), the effective performance between the automatic feature engineering methods (feature extraction, feature selection, and classification) based on DNN from the enriched feature pool (CWT + DCNA), handcraft-method-aided shallow neural networks (SFE + GA + KNN, SFE + SVM), and the effect of enriching feature pool configuration methods (CWT and STHT).

To evaluate the proposed method against the reference methods, the overall classification accuracy (R_f) for each framework has been calculated using Equation (5.10).

$$R_f = \frac{\sum TP}{\sum TS} . 100\% \qquad (5.10)$$

where, $\sum TP$ denotes the summation of the true positives, and $\sum TS$ refers to the total number of samples used in the classifying process. Each framework was executed several times to achieve the average results of overall classifying accuracies for seven defect types. The classification results of all frameworks through two scenarios are shown in Table 5.6.

Table 5.6. The overall identification accuracies of the compared frameworks through two scenarios.

The scenarios	The experiment	The Average Classification Accuracies of Frameworks (%)					
		Fw1	Fw2	Fw3	Fw4	Fw5	Proposed Fw
Scenario 1	Experiment 0	62.18	65.13	54.71	83.50	91.68	100
Scenario 2	Experiment 1	53.43	57.35	72.54	86.65	88.82	100
	Experiment 2	45.31	51.43	68.78	81.51	86.85	100
	Experiment 3	57.62	67.71	74.19	88.30	90.21	100
	Experiment 4	50.17	58.69	72.70	85.90	89.49	100

As can be seen from Table 5.6, the LADT approach performed denoising better than the ANR-GRS method in the three frameworks: Fw1, Fw2, Fw3; however, the identification accuracy results were lesser than the proposed method from 54.69% to 25.81% due to the limitations of those frameworks in engaging with handcraft feature extraction and shallow learning networks. The different results (from 12.7% to 18.49%) between Fw4 and the proposed framework in this paper confirm the high improvement in denoising delicacy of LADT. The

Fw5 results (from 13.15% to 8.32% as lower) demonstrate that the wavelet-based vibration imaging to configure the wealth feature pool achieved a better performance than using STFT. By comparative analysis, the applied framework in this paper outperformed the defect type identification for an MDTF gearbox under variable speeds condition in comparison with those state-of-the-art frameworks, yielding an average classification performance of 100% during two scenarios.

To establish an accurate fault identification framework, an effective denoising technique for the complex gearbox vibration signals is critically needed. The disturbance noises in the vibration signals make the subsequent processes of feature engineering and classification less effective. Therefore, this paper combined LADT for highly effective denoising, VWI for wealthy visual feature pool configuration, and DCNA for high dimensional and automated feature extraction, feature-optimizing selection and classification, and to draw the accurate and stable fault identification framework for an MDTF gearbox under inconsistent speed conditions. Through analysis and experimentation, our proposed methodology achieved the highest classification result, verifying the effectiveness of the proposed model.

5.6 Conclusions

This paper proposed an accurate and stable fault diagnosis framework for multi-degree tooth faults in the gearbox under variable speed conditions. The raw vibration signal obtained from the gearbox contains fault-related information and background noises. To obtain information related to multi-degree tooth faults from the vibration signal, the proposed method preprocesses the raw vibration signal by using the newly developed localized adaptive denoising technique. The localized adaptive denoising technique results in optimized vibration sub-bands with reduced noise. To obtain fault-related information in the form of a time–frequency scale image, a wavelet-based vibration imaging approach is applied to the denoised vibration signal. Finally, these wavelet-based vibration images are provided as an input to a deep convolutional neural network model for fault classification. The deep convolutional neural network is specifically developed for fault diagnosis purposes. To verify the effectiveness of the proposed method, the proposed method was applied to two different datasets. The first dataset had a fixed speed; however, the second dataset consisted of variable speed conditions. On both datasets, the proposed method outperformed the existing state-of-

the-art methods with an average classification accuracy of 100%. In the future, the goal is to apply the proposed fault diagnosis technique to the fault diagnosis of complex rotating machinery, such as centrifugal pumps.

Chapter 6

Summary of Contributions and Future Work

6.1 Introduction

This chapter describes the summation of the main contributions of this thesis and the discussion of the possibilities for future study. The contributions are summarized in Section 6.2, whereas the vision of further research in the future is illustrated in Section 6.3.

6.2 Summary of Contributions

In this thesis, the research works have been conducted to address the difficult and challenging problems appertained to noise reducing techniques of an enhanced VSA, which is an initial-important process to construct a rotation machine conditional monitoring system. From analyzing the vibration characteristic of a gearbox system, the adaptive denoising methodologies have been developed to process nonlinear, nonstationary, and noisy gearbox vibration signals for reducing noise significantly and conserving the fault-related information. Those proposed techniques are well-applicable to construct the sensitive and reliable fault diagnosis models for a multi-level tooth cut gearbox under varying rotational speed conditions. Accordingly, Chapters 2 to 4 explain the procedure to build up the ANR-GRS methodology and its effectiveness in the proposed gearbox fault diagnosis while chapter 5 is about the LADT approach and its application for an accurate and stable fault identification model. The contributions in this dissertation for enhanced vibrations signal processing can be summarized below:

- A reliable fault diagnosis method for a gearbox system with varying rotational speeds was described in Chapter 2. The Proposed method was established for identifying four

fault states of an MTCG gearbox when it operates in varying rotational speed conditions. Adaptive noise reducer based Gaussian reference signal (ANR-GRS) technique was proposed for optimizing gearbox vibration signals in terms of denoising and highlight fault informative elements. ANR-GRS uses the noise-simulated reference signal to access the space between two consecutive sideband frequencies of a vibration spectrum adaptively for removing noise. Next, the machine learning based classification method (OAOMCSVM) was applied to identify those health states of a gearbox using the statistical parameters, which were manually extracted from optimized vibration subbands, as input data. Thus, when those parameters were fed to the support vector machine classifier, the satisfactory accuracy of the classification result was achieved. This model was performed training and testing with the alternative speed-related vibration data for evaluating its reliable capability. Therefore, by the effectiveness of the novel noise reduction technique (ANR-GRS) for vibrations signals, this fault diagnosis framework verified the reliable performance and was well-applicable for diagnosing four fault states of an MTCG gearbox in varying speed conditions.

- Another demand in gearbox fault diagnosis is to investigate gear defects when there exist failures in multiple levels. The fault diagnosis mechanism in Chapter 3, combines the proposed effective noise reduction method ANR-GRS and optimal feature pool configuration based genetic algorithm for identifying six levels of tooth defects (six fault states) of an MTCG gearbox. The manual extraction method tried to extract many signal features in time and frequency domains from the optimized vibration subbands, which were obtained by the ANR-GRS approach and carried the fault information of six fault states, to configure a feature pool. ANR-GRS method is capable of reducing noise significantly and simultaneously keeping the fault-related ingredients in vibration signals of an MTCG gearbox. Then, the genetic algorithm (GA) was applied to perform a heuristic searching strategy in the extracted feature pool to select the most discriminative features for the configuration of an optimal feature pool. Thus, a simple classifying algorithm such as k-nearest neighbor was used for identifying fault states into the respective categories using the optimal subset outputted from a genetic algorithm based feature selection. Then, the combination model (ANR-GRS and GA)

yielded the superior performance result to identify the defects of multiple levels. The conclusion could be clarified that the proposed fault identification model can be well-applied for discriminate fault states of an MTCG gearbox system.

- In Chapter 4, the sensitive and speed invariant gearbox fault diagnosis scheme was constructed by incorporating and utilizing the advantages of the novel adaptive noise control (ANC) technique and a deep neural network. The vibration data collected from seven fault types of multi-level tooth cut faults under varying rotational speeds, which were observed in this chapter, are very heterogeneous. In addition to the aforementioned attributes as: nonlinear, nonstationary, and noisy, the vibration signals were to vibrate of similar behaviors of multi-level tooth faults. Thus, the process to highlight the representative features of each fault type for separating the fault types was much arduous. This incorporated model addressed those difficult problems. Firstly, ANC performed an optimization process of raw vibration signals to output the optimal vibration subbands, which comprise the plurality of fault-related frequency components. Next, a deep learning model as stacked sparse autoencoder-based deep neural network (SSA-DNN), which was established by stacking several sparse autoencoders as the network hidden layers and SoftMax activation function for output classification layer, was used for feature engineering and classification. Accordingly, SSA-DNN performed feature extraction automatically to numerous latent representation features, from the optimal vibration subbands, then the deep network architecture performed the fine-tuning process to select the most discriminating fault features based on a back-propagation algorithm and reconstruction error minimization. Those fault types representation features were highly-distinguishable and could be identified easily via the output SoftMax layer of the network architecture. Therefore, the proposed combination framework provided excellent diagnosis results and can be applied for various fault types of a gearbox under variable rotational speed conditions.
- In Chapter 5, the novel localized adaptive denoising technique (LADT) was proposed. Then, the stable and accurate gearbox fault diagnosis scheme for multi-degree gear faults under inconsistent shaft speed conditions was built up by combining LADT, wavelet vibration imaging (WVI), and deep convolutional neural network architecture

(DCNA). The objective of this study is to develop more denoising techniques. Because of the dissimilarity between principle frequency segments (PFS) in a frequency spectrum of a vibration signal, for reducing more noise components, LADT was constructed by applying the ANR-GRS method on each PFS to figure out the localized optimal parameter. As the result, LADT could reduce noise better than ANR-GRS, and preserve original fault useful information, it was well-applicable in company with feature engineering and classification method for establishing a gearbox fault diagnosis model. Then a visualized feature pool configuration was formed by using wavelet based vibration imaging methods for optimized vibration subbands outputted from LADT. The scalogram images samples related to fault states were provided to DCNA for classification, which yielded satisfactory identification results. The efficiency of the reliable and accurate gearbox fault diagnosis framework verified the superiority of LADT for noise reduction, accurate and reliable performance, and qualification for identifying many fault types of multi-degree tooth fault under varying rotational speed conditions.

6.3 Future Work

As described in the previous chapters the main focus of this dissertation is to develop denoising techniques for condition monitoring of the industrial machinery. Although the effectiveness of the proposed noise reduction techniques (ANR-GRS and LADT) have outperformed the conventional methodologies, there are numerous demands related to vibration characteristic of the observed industrial machines as; bearing, centrifugal pumps, induction motors, pipelines … Thus those denoising methods are needed to improve for compliance. Moreover, condition monitoring fault diagnosis models are commonly constructed by combining signal processing, feature engineering, and machine learning based classification. The efficient denoising techniques and the deep neural network-based approaches are demonstrated the superiority performance for the establishment of fault identification models for multiple fault types under variable speeds conditions. However, there exist many difficult problems of condition monitoring models for different industrial machines, those need to develop and improve more in future works. Some of these further research topics are listed below:

- The adaptive noise reducing methods should be developed more in future research for improving capabilities of noise reduction, online or real-time health monitoring. Accordingly, the adaptive algorithms in the proposed denoising models can be improved by examining other methods as fuzzy or machine learning based algorithms for replacement. Moreover, the data acquisition and signal processing systems should be communicated to the controlling systems of rotation machines for early fault detection and maintenance. For establishing online diagnosis and prognosis systems for rotation machines, the denoising techniques should be further investigated in practice for real time processing.

- A possible and fascinating performance for future study of the works described in this thesis is a data fusion technique constructed by simultaneous monitoring of several sensors. Data fusion can be used to provide more fault-related information by accumulating data from each sensor channel. Indeed, the efficient denoising methods and data fusion approaches can be applied for fetching discriminative features, which are useful for the classification process of a fault diagnosis system. Therefore, data fuse by using multiple sensors is a practical concentration for better performance of condition monitoring of rotation machines.

- Transfer learning is an advanced technique of machine learning and can be well-applicable for fault identification of further researches. The transfer learning method focuses on knowledge-based performance, the knowledge gathered by solving a problem and using it for related problems. In that way, solution knowledge can be applied for different objects whose characteristics are analogous. In transfer learning, the properties of the source task are discovered and transferred to the target task (e.g. mapping function, weights…) for expanding the learning performance of the target from the knowledge of the source. By using transfer learning, the learning speed can be improved and the amount of required data can be reduced. Those are superiorities of transfer learning performance and can be applied for designing the real-time or online condition monitoring systems of rotation machines.

Publications

International Journal Papers

1. **C. D. Nguyen**, A. Prosvirin, and J.-M. Kim, "A Reliable Fault Diagnosis Method for a Gearbox System with Varying Rotational Speeds", *Sensors*, vol. 20, no. 11, 2020.

2. **C. D. Nguyen**, A. Prosvirin, and J.-M. Kim, "Construction of a Sensitive and Speed Invariant Gearbox Fault Diagnosis Model Using an Incorporated Utilizing Adaptive Noise Control and a Stacked Sparse Autoencoder-Based Deep Neural Network," *Sensors*, vol. 21, no. 18, 2021

3. **C. D. Nguyen**, Z. Ahmad and J.-M. Kim, "Gearbox Fault Identification Framework Based on Novel Localized Adaptive Denoising Technique, Wavelet-Based Vibration Imaging, and Deep Convolutional Neural Network," *Appl.*,vol. 11, no. 7575, 2021

Book Chapters

1. **C. D. Nguyen**, A. Prosvirin, and J.-M. Kim, "Fault Identification of Multi-level Gear Defects Using Adaptive Noise Control and a Genetic Algorithm," In: Singh M., Kang DK., Lee JH., Tiwary U.S., Singh D., Chung WY. (eds) Intelligent Human Computer Interaction. IHCI 2020. Lecture Notes in Computer Science, vol 12615. Springer.

International Conferences

4. **C. D. Nguyen**, A. Prosvirin, and J.-M. Kim, "Fault Identification of Multi-level Gear Defects Using Adaptive Noise Control and a Genetic Algorithm," in *12th International Conference on Intelligent Human Computer Interaction, IHCI2020*, 2020.

Domestic Conferences

1. **C. D. Nguyen** and J.-M. Kim, " Gearbox Fault Diagnosis Based on Adaptive De-noising and Support Vector Machine," in The Engineering and Arts Society in Korea, University of Ulsan, 2019.

References

[1] X. Jin, F. Cheng, Y. Peng, W. Qiao, and L. Qu, "Drivetrain gearbox fault diagnosis: Vibration-and current-based approaches," *IEEE Ind. Appl. Mag.*, vol. 24, no. 6, pp. 56–66, Nov. 2018.

[2] P. D. Samuel, D. J. Pines, and D. G. Lewicki, "Comparison of stationary and non-stationary metrics for detecting faults in helicopter gearboxes," *J. Am. Helicopter Soc.*, vol. 45, no. 2, pp. 125–136, Apr. 2000.

[3] J. W. Baxter and J. R. Bumby, "An explanation for the asymmetry of the modulation sidebands about the tooth meshing frequency in epicyclic gear vibration," *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, vol. 199, no. 1, pp. 65–70, 1985.

[4] R. Errichello, "How to analyze gear failures," *J. Fail. Anal. Prev.*, vol. 2, no. 6, pp. 8–16, 2002.

[5] L. E. Alban, "Failures of Gears," in *Failure Analysis and Prevention*, vol. 11, R. J. S. William T. Becker, Ed. ASM International, 2002.

[6] L. Wang, J. Chu, and J. Wu, "Selection of optimum maintenance strategies based on a fuzzy analytic hierarchy process," *Int. J. Prod. Econ.*, vol. 107, no. 1, pp. 151–163, May 2007.

[7] R. Ahmad and S. Kamaruddin, "An overview of time-based and condition-based maintenance in industrial application," *Comput. Ind. Eng.*, vol. 63, no. 1, pp. 135–149, Aug. 2012.

[8] D. McMillan and G. W. Ault, "Quantification of condition monitoring benefit for offshore wind turbines," *Wind Eng.*, vol. 31, no. 4, pp. 267–285, May 2007.

[9] J. Van Dam and L. J. Bond, "Economics of online structural health monitoring of wind turbines: Cost benefit analysis," in *AIP Conference Proceedings*, 2015, vol. 1650, no. 1, pp. 899–908.

[10] N. K. Verma, S. Khatravath, and A. Salour, "Cost benefit analysis for condition based maintenance," in *PHM 2013 - 2013 IEEE International Conference on Prognostics and Health Management, Conference Proceedings*, 2013.

[11] C. Malla and I. Panigrahi, "Review of Condition Monitoring of Rolling Element Bearing Using Vibration Analysis and Other Techniques," *J. Vib. Eng. Technol.*, vol. 7, no. 4, pp. 407–414, Aug. 2019.

[12] R. B. Randall, *Vibration-Based Condition Monitoring: Industrial, Aerospace and Automotive Applications*. wiley, 2010.

[13] A. R. Al-Obaidi and H. Towsyfyan, "An experimental study on vibration signatures for detecting incipient cavitation in centrifugal pumps based on envelope spectrum analysis," *J. Appl. Fluid Mech.*, vol. 12, no. 6, pp. 2057–2067, 2019.

[14] A. S. Sait and Y. I. Sharaf-Eldeen, "A review of gearbox condition monitoring based on vibration analysis techniques diagnostics and prognostics," in *Conference Proceedings of the Society for Experimental Mechanics Series*, 2011, vol. 5, pp. 307–324.

[15] J. Mączak and M. Jasiński, "Model-based detection of local defects in gears," *Arch. Appl. Mech.*, vol. 88, no. 1–2, pp. 215–231, Feb. 2018.

[16] R. R. B., "Cepstrum Analysis and Gearbox Fault Diagnosis," *Maint. Manag. Int.*, vol. 3, no. 3, pp. 183–208, 1982.

[17] P. D. McFadden, "Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration," *J. Vib. Acoust. Trans. ASME*, vol. 108, no. 2, pp. 165–170, 1986.

[18] W. Bartelmus, F. Chaari, R. Zimroz, and M. Haddar, "Modelling of gearbox dynamics under time-varying nonstationary load for distributed fault detection and diagnosis," *Eur. J. Mech. A/Solids*, vol. 29, no. 4, pp. 637–646, Jul. 2010.

[19] N. Sawalhi, "Vibration Sideband Modulations and Harmonics Separation of a Planetary Helicopter Gearbox with Two Different Configurations," *Adv. Acoust. Vib.*, vol. 2016, 2016.

[20] Z. Chen and Y. Shao, "Dynamic simulation of spur gear with tooth root crack propagating along tooth width and crack depth," *Eng. Fail. Anal.*, vol. 18, no. 8, pp. 2149–2164, Dec. 2011.

[21] F. Chaari, T. Fakhfakh, and M. Haddar, "Analytical modelling of spur gear tooth crack and influence on gearmesh stiffness," *Eur. J. Mech. A/Solids*, vol. 28, no. 3, pp. 461–468, May 2009.

[22] C. Cempel and M. Tabaszewski, "Multidimensional condition monitoring of machines in non-stationary operation," *Mech. Syst. Signal Process.*, vol. 21, no. 3, pp. 1233–1241, Apr. 2007.

[23] D. R. Houser, "Gear Noise and Vibration Prediction and Control Methods," in *Handbook of Noise and Vibration Control*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2008, pp. 847–856.

[24] T. Reimann, T. Herzog, D. Kadach, and K. Stahl, "The influence of friction on the tooth normal force of spur and helical gears," *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, vol. 233, no. 21–22, pp. 7391–7400, Nov. 2019.

[25] Z. Zhong, Z. Jiang, Y. Long, and X. Zhan, "Analysis on the noise for the different gearboxes of the heavy truck," *Shock Vib.*, vol. 2015, 2015.

[26] S. Wu, M. J. Zuo, and A. Parey, "Simulation of spur gear dynamics and estimation of fault growth," *J. Sound Vib.*, vol. 317, no. 3–5, pp. 608–624, Nov. 2008.

[27] P. D. Samuel and D. J. Pines, "A review of vibration-based techniques for helicopter transmission diagnostics," *J. Sound Vib.*, vol. 282, no. 1–2, pp. 475–508, Apr. 2005.

[28] T. Barszcz and R. B. Randall, "Application of spectral kurtosis for detection of a tooth crack in the planetary gear of a wind turbine," *Mech. Syst. Signal Process.*, vol. 23, no. 4, pp. 1352–1365, May 2009.

[29] S. Braun, "The synchronous (time domain) average revisited," *Mech. Syst. Signal Process.*, vol. 25, no. 4, pp. 1087–1102, May 2011.

[30] X. Fan and M. J. Zuo, "Gearbox fault feature extraction using Hilbert transform, S-

transform, and a statistical indicator," *J. Test. Eval.*, vol. 35, no. 5, pp. 477–485, Sep. 2007.

[31]   H. Li, L. Fu, and Z. Shi, "Research on the gear fault diagnosis using order envelope spectrum," in *2009 International Conference on Measuring Technology and Mechatronics Automation, ICMTMA 2009*, 2009, vol. 1, pp. 711–714.

[32]   R. B. Randall, "Basic Signal Processing Techniques," in *Vibration-Based Condition Monitoring*, wiley, 2011, pp. 63–141.

[33]   A. G. Piersol and T. L. Paez, *Harris' Shock and Vibration Handbook*, 6th ed. McGraw-Hill New York, 2010.

[34]   P. D. McFadden, "Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration," *J. Vib. Acoust. Trans. ASME*, vol. 108, no. 2, pp. 165–170, Apr. 1986.

[35]   F. Combet, L. Gelman, P. Anuzis, and R. Slater, "Vibration detection of local gear damage by advanced demodulation and residual techniques," *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.*, vol. 223, no. 5, pp. 507–514, Aug. 2009.

[36]   G. Dalpiaz, A. Rivola, and R. Rubini, "Effectiveness and sensitivity of vibration processing techniques for local fault detection in gears," *Mech. Syst. Signal Process.*, vol. 14, no. 3, pp. 387–412, May 2000.

[37]   I. Howard, S. Jia, and J. Wang, "The dynamic modelling of a spur gear in mesh including friction and a crack," *Mech. Syst. Signal Process.*, vol. 15, no. 5, pp. 831–853, Sep. 2001.

[38]   R. M. Stewart, *Some Useful Data Analysis Techniques for Gearbox Diagnostics - R. M. Stewart - Google Books*. University of Southampton, 1977.

[39]   R. B. Randall, N. Sawalhi, and M. Coats, "A comparison of methods for separation of deterministic and random signals," *Int. J. Cond. Monit.*, vol. 1, no. 1, pp. 11–19, Nov. 2011.

[40]   Z. Tian, M. J. Zuo, and S. Wu, "Crack propagation assessment for spur gears using model-based analysis and simulation," *J. Intell. Manuf.*, vol. 23, no. 2, pp. 239–253, Apr. 2012.

[41]   A. Wong, W. Wang, and A. K. Wong, "Autoregressive model-based gear fault diagnosis," *J. Vib. Acoust.*, 2002.

[42]   X. Zhou, Y. Shao, Y. Lei, and M. Zuo, "Time-varying meshing stiffness calculation and vibration analysis for a 16DOF dynamic model with linear crack growth in a pinion," *J. Vib. Acoust.*, vol. 134, no. 1, 2012.

[43]   W. J. Wang and P. D. McFadden, "Early detection of gear failure by vibration analysis—I. calculation of the time-frequency distribution," *Mech. Syst. Signal Process.*, vol. 7, no. 3, pp. 193–203, May 1993.

[44]   W. J. Staszewski, K. Worden, and G. R. Tomlinson, "Time-frequency analysis in gearbox fault detection using the Wigner-Ville distribution and pattern recognition," *Mech. Syst. Signal Process.*, vol. 11, no. 5, pp. 673–692, Sep. 1997.

[45] Z. K. Peng and F. L. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: A review with bibliography," *Mech. Syst. Signal Process.*, vol. 18, no. 2, pp. 199–221, Mar. 2004.

[46] M. Lokesha, M. Majumder, K. Ramachandran, and K. Raheem, "Fault diagnosis in gear using wavelet envelope power spectrum," *Int. J. Eng. Sci. Technol.*, vol. 3, no. 8, pp. 156–167, Jan. 1970.

[47] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, 1998.

[48] J. Cai and X. Li, "Gear Fault Diagnosis Based on Empirical Mode Decomposition and 1.5 Dimension Spectrum," *Shock Vib.*, vol. 2016, 2016.

[49] J. Wang, J. Li, H. Wang, and L. Guo, "Composite fault diagnosis of gearbox based on empirical mode decomposition and improved variational mode decomposition," *J. Low Freq. Noise Vib. Act. Control*, vol. 40, no. 1, pp. 332–346, Mar. 2021.

[50] V. Puliafito, S. Vergura, and M. Carpentieri, "Fourier, wavelet, and Hilbert-Huang transforms for studying electrical users in the time and frequency domain," *Energies*, vol. 10, no. 1, 2017.

[51] Y. Zhang, B. Tang, and X. Xiao, "Time-frequency interpretation of multi-frequency signal from rotating machinery using an improved Hilbert-Huang transform," *J. Int. Meas. Confed.*, vol. 82, pp. 221–239, Mar. 2016.

[52] S. X. Yang, J. S. Hu, Z. T. Wu, and G. B. Yan, "Comparison of vibration signals time-frequency analysis between EMD-based HT and STFT method in rotating machinery," in *Proceedings of the Chinese Society of Electrical Engineering*, 2003, pp. 102–107.

[53] Q. Yang and D. An, "EMD and Wavelet Transform Based Fault Diagnosis for Wind Turbine Gear Box," *Adv. Mech. Eng.*, vol. 5, p. 212836, Jan. 2013.

[54] M. Van, H. J. Kang, and K. S. hin, "Rolling element bearing fault diagnosis based on non-local means de-noising and empirical mode decomposition," *IET Sci. Meas. Technol.*, vol. 8, no. 6, pp. 571–578, Nov. 2014.

[55] Y. Guo, Q. N. Liu, X. Wu, and J. Na, "Gear fault diagnosis based on narrowband demodulation with frequency shift and spectrum edit," *Int. J. Eng. Technol. Innov.*, vol. 6, no. 4, pp. 243–254, 2016.

[56] W. Li, J. Zhang, L. Cui, L. Gao, and F. Zhang, "Application of multiwavelet adaptive threshold denoising in the fault diagnosis of gearbox," in *Proceedings - 3rd International Conference on Measuring Technology and Mechatronics Automation, ICMTMA 2011*, 2011, vol. 1, pp. 551–554.

[57] J. Chen, Y. Zi, Z. He, and X. Wang, "Adaptive redundant multiwavelet denoising with improved neighboring coefficients for gearbox fault detection," *Mech. Syst. Signal Process.*, vol. 38, no. 2, pp. 549–568, Jul. 2013.

[58] T. W. Rauber, F. De Assis Boldt, and F. M. Varejão, "Heterogeneous feature models and feature selection applied to bearing fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 637–646, Jan. 2015.

[59] P. Li, Q. He, and F. Kong, "A new method of feature extraction for gearbox vibration signals," in *Proceedings - 2010 3rd International Congress on Image and Signal Processing, CISP 2010*, 2010, vol. 8, pp. 3980–3984.

[60] X. Chen, C. Yin, and W. He, "Feature extraction of gearbox vibration signals based on EEMD and sample entropy," in *Proceedings - 2013 10th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2013*, 2013, pp. 811–815.

[61] J. Wodecki, P. Stefaniak, J. Obuchowski, A. Wylomanska, and R. Zimroz, "Combination of principal component analysis and time-frequency representations of multichannel vibration data for gearbox fault detection," *J. Vibroengineering*, vol. 18, no. 4, pp. 2167–2175, 2016.

[62] W. Li, L. Zhang, and Y. Xu, "Gearbox pitting detection using linear discriminant analysis and distance preserving self-organizing map," in *2012 IEEE I2MTC - International Instrumentation and Measurement Technology Conference, Proceedings*, 2012, pp. 2225–2229.

[63] W. H. Li, T. L. Shi, and S. Z. Yang, "An approach for mechanical fault classification based on generalized discriminant analysis," *Front. Mech. Eng. China*, vol. 1, no. 3, pp. 292–298, Sep. 2006.

[64] X. Tian, J. Lin, K. R. Fyfe, and M. J. Zuo, "Gearbox fault diagnosis using independent component analysis in the frequency domain and wavelet filtering," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2003, vol. 2, pp. 245–248.

[65] Z. Yang, W. I. Hoi, and J. Zhong, "Gearbox fault diagnosis based on artificial neural network and genetic algorithms," in *Proceedings 2011 International Conference on System Science and Engineering, ICSSE 2011*, 2011, pp. 37–42.

[66] Y. Lei and M. J. Zuo, "Gear crack level identification based on weighted K nearest neighbor classification algorithm," *Mech. Syst. Signal Process.*, vol. 23, no. 5, pp. 1535–1547, 2009.

[67] B. Samanta, "Gear fault detection using artificial neural networks and support vector machines with genetic algorithms," *Mech. Syst. Signal Process.*, vol. 18, no. 3, pp. 625–644, May 2004.

[68] P. G. Sreenath, G. Praveen Kumare, S. Pravin, K. N. Vikram, and M. Saimurugan, "Automobile Gearbox Fault Diagnosis Using Naive Bayes and Decision Tree Algorithm," *Appl. Mech. Mater.*, vol. 813–814, pp. 943–948, Nov. 2015.

[69] D. Yang, Y. Liu, S. Li, X. Li, and L. Ma, "Gear fault diagnosis based on support vector machine optimized by artificial bee colony algorithm," *Mech. Mach. Theory*, vol. 90, pp. 219–229, Aug. 2015.

[70] S. Afaq Ali Shah, M. Bennamoun, and F. Boussaid, "Iterative deep learning for image set based face and object recognition," *Neurocomputing*, vol. 174, pp. 866–874, Jan. 2016.

[71] R. Sarikaya, G. E. Hinton, and A. Deoras, "Application of deep belief networks for natural language understanding," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 22,

no. 4, pp. 778–784, Apr. 2014.

[72] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, Apr. 2016.

[73] Y. Zhang, W. Chan, and N. Jaitly, "Very deep convolutional networks for end-to-end speech recognition," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2017.

[74] C. C. Aggarwal, *Neural Networks and Deep Learning*. Springer, 2018.

[75] S. R. Saufi, Z. A. Bin Ahmad, M. S. Leong, and M. H. Lim, "Low-Speed Bearing Fault Diagnosis Based on ArSSAE Model Using Acoustic Emission and Vibration Signals," *IEEE Access*, vol. 7, pp. 46885–46897, 2019.

[76] M. Sohaib and J. M. Kim, "Reliable Fault Diagnosis of Rotary Machine Bearings Using a Stacked Sparse Autoencoder-Based Deep Neural Network," *Shock Vib.*, vol. 2018, 2018.

[77] S. R. Saufi, Z. A. Bin Ahmad, M. S. Leong, and M. H. Lim, "Challenges and opportunities of deep learning models for machinery fault detection and diagnosis: A review," *IEEE Access*, vol. 7, pp. 122644–122662, 2019.

[78] L. Jing, M. Zhao, P. Li, and X. Xu, "A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox," *Meas. J. Int. Meas. Confed.*, vol. 111, pp. 1–10, Dec. 2017.

[79] S. R. Saufi, Z. A. Bin Ahmad, M. S. Leong, and M. H. Lim, "Gearbox Fault Diagnosis Using a Deep Learning Model with Limited Data Sample," *IEEE Trans. Ind. Informatics*, vol. 16, no. 10, pp. 6263–6271, 2020.

[80] J. McNames, "Fourier series analysis of epicyclic gearbox vibration," *J. Vib. Acoust. Trans. ASME*, vol. 124, no. 1, pp. 150–160, 2002.

[81] T. Fakhfakh, F. Chaari, and M. Haddar, "Numerical and experimental analysis of a gear system with teeth defects," *Int. J. Adv. Manuf. Technol.*, vol. 25, no. 5–6, pp. 542–550, 2005.

[82] F. Chaari, W. Bartelmus, R. Zimroz, T. Fakhfakh, and M. Haddar, "Gearbox vibration signal amplitude and frequency modulation," *Shock Vib.*, vol. 19, no. 4, pp. 635–652, 2012.

[83] J. S. Mitchell, "An Introduction to Machinery Analysis and Monitoring," *Comput. Eng.*, vol. 10, no. 12, pp. 314–315, 1991.

[84] S. B. Ghodake, P. A. K. Mishra, and P. A. V Deokar, "A Review on Fault Diagnosis of Gear-Box by Using Vibration Analysis Method," *IPASJ Int. J. Mech. Eng.*, vol. 4, no. 1, pp. 31–35, 2016.

[85] K. U. S. Chetan Ramesh Patil, Prasad Prabhakar Kulkarni, Nitin Narayan Sarode, "Gearbox Noise & Vibration Prediction and Control," *Int. Res. J. Eng. Technol.*, vol. 04, no. 2, pp. 873–877, 2017.

[86] A. K. Boyat and B. K. Joshi, "A Review Paper : Noise Models in Digital Image Processing," *Signal Image Process. An Int. J.*, vol. 6, no. 2, pp. 63–75, 2015.

[87] R. B. Randall, *Frequency Analysis*, 3rd ed. Nairobi, Denmark: Brüel & Kjaer, 1987.

[88] G. Matz and F. Hlawatsch, "Wigner distributions (nearly) everywhere: Time-frequency analysis of signals, systems, random processes, signal spaces, and frames," *Signal Processing*, vol. 83, no. 7, pp. 1355–1378, 2003.

[89] W. Zhang, H. Wang, R. Teng, and S. Xu, "Application of rank-order morphological filter in vibration signal de-noising," in *Proceedings - 2010 3rd International Congress on Image and Signal Processing, CISP 2010*, 2010, vol. 8, pp. 4025–4027.

[90] J. Antoni, "Cyclostationarity by examples," *Mech. Syst. Signal Process.*, vol. 23, no. 4, pp. 987–1036, 2009.

[91] R. B. Randall, J. Antoni, and S. Chobsaard, "The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals," *Mech. Syst. Signal Process.*, vol. 15, no. 5, pp. 945–962, 2001.

[92] W. Touti, M. Salah, S. Ben Salem, K. Bacha, and A. Chaari, "Spur gearbox mixed fault detection using vibration envelope and motor stator current signatures analysis," in *2016 17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering, STA 2016 - Proceedings*, 2017, pp. 193–198.

[93] R. Jiang, S. Liu, Y. Tang, and Y. Liu, "A novel method of fault diagnosis for rolling element bearings based on the accumulated envelope spectrum of the wavelet packet," *JVC/Journal Vib. Control*, vol. 21, no. 8, pp. 1580–1593, 2015.

[94] H. Xie, J. Lin, Y. Lei, and Y. Liao, "Fast-varying AM-FM components extraction based on an adaptive STFT," *Digit. Signal Process. A Rev. J.*, vol. 22, no. 4, pp. 664–670, 2012.

[95] Alan V Oppenheim and Ronald W Schafer, *Discrete-time signal processing*. Englewood Cliffs, NJ, USA: Prentice Hall, 1989.

[96] Martin Vetterli and Jelena Kovačević, *Wavelets and Subband Coding*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1995.

[97] J. Olkkonen, *Discrete Wavelet Transforms Theory and Applications*. Intech, 2011.

[98] J. Liu, "Shannon wavelet spectrum analysis on truncated vibration signals for machine incipient fault detection," *Meas. Sci. Technol.*, vol. 23, no. 5, p. 055604, Apr. 2012.

[99] K. Aharamuthu and E. P. Ayyasamy, "Application of discrete wavelet transform and Zhao-Atlas-Marks transforms in non stationary gear fault diagnosis," *J. Mech. Sci. Technol.*, vol. 27, no. 3, pp. 641–647, Mar. 2013.

[100] M. Kang, J. Kim, J. M. Kim, A. C. C. Tan, E. Y. Kim, and B. K. Choi, "Reliable fault diagnosis for low-speed bearings using individually trained support vector machines with kernel discriminative feature analysis," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2786–2797, May 2015.

[101] S. J. Loutridis, "Damage detection in gear systems using empirical mode decomposition," *Eng. Struct.*, vol. 26, no. 12, pp. 1833–1841, Oct. 2004.

[102] C. Zhang, Z. Peng, S. Chen, Z. Li, and J. Wang, "A gearbox fault diagnosis method

based on frequency-modulated empirical mode decomposition and support vector machine," *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, vol. 232, no. 2, pp. 369–380, 2018.

[103] K. Chen, X. C. Zhou, J. Q. Fang, P. F. Zheng, and J. Wang, "Fault Feature Extraction and Diagnosis of Gearbox Based on EEMD and Deep Briefs Network," *Int. J. Rotating Mach.*, vol. 2017, 2017.

[104] M. Buzzoni, E. Mucchi, G. D'Elia, and G. Dalpiaz, "Diagnosis of Localized Faults in Multistage Gearboxes: A Vibrational Approach by Means of Automatic EMD-Based Algorithm," *Shock Vib.*, vol. 2017, 2017.

[105] B. Liu, S. Riemenschneider, and Y. Xu, "Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum," *Mech. Syst. Signal Process.*, vol. 20, no. 3, pp. 718–734, Apr. 2006.

[106] A. Y. Goharrizi and N. Sepehri, "Internal leakage detection in hydraulic actuators using empirical mode decomposition and hilbert spectrum," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 2, pp. 368–378, Feb. 2012.

[107] G. D. Han, S. T. Wan, Z. J. Lv, R. H. Liu, J. Wang, and G. J. Tang, "The analysis of gearbox fault diagnosis research based on the EMD and hilbert envelope demodulation," in *Advanced Materials Research*, 2014, vol. 926–930, pp. 1800–1805.

[108] Z. K. Peng, P. W. Tse, and F. L. Chu, "A comparison study of improved Hilbert-Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing," *Mech. Syst. Signal Process.*, vol. 19, no. 5, pp. 974–988, 2005.

[109] J. Lin, C. Dou, and Q. Wang, "Comparisons of MFDFA, EMD and WT by neural network, Mahalanobis distance and SVM in fault diagnosis of gearboxes," *Sound Vib.*, vol. 52, no. 2, pp. 11–15, Apr. 2018.

[110] A. H. Zamanian and A. Ohadi, "Gear fault diagnosis based on Gaussian correlation of vibrations signals and wavelet coefficients," *Appl. Soft Comput. J.*, vol. 11, no. 8, pp. 4807–4819, Dec. 2011.

[111] A. S. Raj and N. Murali, "Morlet wavelet UDWT denoising and EMD based bearing fault diagnosis," *Electronics*, vol. 17, no. 1, pp. 1–8, Jun. 2013.

[112] X. Wang, Y. Zi, and Z. He, "Multiwavelet denoising with improved neighboring coefficients for application on rolling bearing fault diagnosis," *Mech. Syst. Signal Process.*, vol. 25, no. 1, pp. 285–304, 2011.

[113] Simon S Haykin, *Adaptive Filter Theory (4th Edition)*. Upper Saddle River, N.J.: Prentice Hall, 2002.

[114] K. A. Lee, W. S. Gan, and S. M. Kuo, *Subband Adaptive Filtering: Theory and Implementation*. Chichester, UK: John Wiley and Sons, 2009.

[115] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mech. Syst. Signal Process.*, vol. 108, pp. 33–47, Aug. 2018.

[116] J. Liu and E. Zio, "Feature vector regression with efficient hyperparameters tuning and

geometric interpretation," *Neurocomputing*, vol. 218, pp. 411–422, Dec. 2016.

[117] A. Widodo *et al.*, "Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine," *Expert Syst. Appl.*, vol. 36, no. 3 PART 2, pp. 7252–7261, Apr. 2009.

[118] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag New York, Inc., 1995.

[119] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge, UK: Cambridge University Press, 1999.

[120] C. W. Hsu and C. J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Networks*, vol. 13, no. 2, pp. 415–425, Mar. 2002.

[121] X. Fan and M. J. Zuo, "Gearbox fault detection using Hilbert and wavelet packet transform," *Mech. Syst. Signal Process.*, vol. 20, no. 4, pp. 966–982, 2006.

[122] R. S. Figliola and D. E. Beasley, "Sampling, digital devices, and data acquisition," in *Theory and Design for Mechanical Measurements, Fifth Edition*, Hoboken, NJ: John Wiley & Sons, 2011, pp. 260–303.

[123] Widrow B. and Stearns S. D., *Adaptive Signal Processing*. Upper Saddle River, NJ: Prentice Hall, 1985.

[124] W. Caesarendra and T. Tjahjowidodo, "A Review of Feature Extraction Methods in Vibration-Based Condition Monitoring and Its Application for Degradation Trend Estimation of Low-Speed Slew Bearing," *Machines*, vol. 5, no. 4, p. 21, Sep. 2017.

[125] J. D. Rodríguez, A. Pérez, and J. A. Lozano, "Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 569–575, 2010.

[126] C. D. Nguyen, A. Prosvirin, and J. M. Kim, "A reliable fault diagnosis method for a gearbox system with varying rotational speeds," *Sensors (Switzerland)*, vol. 20, no. 11, 2020.

[127] M. A. M. Ariff and B. C. Pal, "Coherency identification in interconnected power system - An independent component analysis approach," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1747–1755, 2013.

[128] V. H. Nguyen and J. C. Golinval, "Fault detection based on Kernel Principal Component Analysis," *Eng. Struct.*, vol. 32, no. 11, pp. 3683–3691, Nov. 2010.

[129] D. G. Ece and M. Baaran, "Condition monitoring of speed controlled induction motors using wavelet packets and discriminant analysis," *Expert Syst. Appl.*, vol. 38, no. 7, pp. 8079–8086, 2011.

[130] 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 Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3299–3310, 2016.

[131] H. Yigit, "A weighting approach for KNN classifier," *2013 Int. Conf. Electron.*

Comput. Comput. ICECCO 2013, vol. 2, no. 2, pp. 228–231, 2013.

[132] W. Bartelmus and R. Zimroz, "A new feature for monitoring the condition of gearboxes in non-stationary operating conditions," *Mech. Syst. Signal Process.*, vol. 23, no. 5, pp. 1528–1534, Jul. 2009.

[133] D. Han, N. Zhao, and P. Shi, "Gear fault feature extraction and diagnosis method under different load excitation based on EMD, PSO-SVM and fractal box dimension," *J. Mech. Sci. Technol.*, vol. 33, no. 2, pp. 487–494, Feb. 2019.

[134] S. S. Ajanalkar, P. V Bute, and G. D. Shrigandhi, "Gear Fault Prediction by using Artificial Neural Network (ANN) – A Review," vol. 7, no. 7, pp. 55–58, 2017.

[135] G. E. H. and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science (80-.).*, vol. 313, no. July, pp. 504–507, 2006.

[136] 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 (Switzerland)*, vol. 17, no. 3, pp. 1–17, 2017.

[137] R. N. Bracewell, Ronald Newbold and Bracewell, *The Fourier Transform & Its Applications*. McGraw-Hill New York, 1986.

[138] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, "Greedy layer-wise training of deep networks," *Adv. Neural Inf. Process. Syst.*, no. 1, pp. 153–160, 2007.

[139] T. Van Erven and P. Harremoës, "Rényi Divergence and Kullback – Leibler Divergence," vol. 60, no. 7, pp. 3797–3820, 2014.

[140] W. Zhao, Z. Wang, C. Lu, J. Ma, and L. Li, "Fault diagnosis for centrifugal pumps using deep learning and softmax regression," *Proc. World Congr. Intell. Control Autom.*, vol. 2016-Septe, pp. 165–169, 2016.

[141] A. Coates, H. Lee, and A. Y. Ng, "An analysis of single-layer networks in unsupervised feature learning," *J. Mach. Learn. Res.*, vol. 15, pp. 215–223, 2011.

[142] J. B. Yu, "Evolutionary manifold regularized stacked denoising autoencoders for gearbox fault diagnosis," *Knowledge-Based Syst.*, vol. 178, pp. 111–122, 2019.

[143] M. Amar, I. Gondal, and C. Wilson, "Vibration spectrum imaging: A novel bearing fault classification approach," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 494–502, 2015.

[144] M. Nie and L. Wang, "Review of condition monitoring and fault diagnosis technologies for wind turbine gearbox," *Procedia CIRP*, vol. 11, no. Cm, pp. 287–290, 2013.

[145] T. Praveenkumar, M. Saimurugan, P. Krishnakumar, and K. I. Ramachandran, "Fault diagnosis of automobile gearbox based on machine learning techniques," in *Procedia Engineering*, 2014, vol. 97, pp. 2092–2098.

[146] D. Goyal, Vanraj, B. S. Pabla, and S. S. Dhami, "Condition Monitoring Parameters for Fault Diagnosis of Fixed Axis Gearbox: A Review," *Arch. Comput. Methods Eng.*, vol. 24, no. 3, pp. 543–556, Jul. 2017.

[147] H. K. Kohler, A. Pratt, and A. M. Thompson, "Dynamics and Noise of Parallel-Axis Gearing," *Proc. Inst. Mech. Eng. Conf. Proc.*, vol. 184, no. 15, pp. 111–121, Sep. 1969.

[148] C. D. Nguyen, A. E. Prosvirin, C. H. Kim, and J.-M. Kim, "Construction of a Sensitive and Speed Invariant Gearbox Fault Diagnosis Model Using an Incorporated Utilizing Adaptive Noise Control and a Stacked Sparse Autoencoder-Based Deep Neural Network," *Sensors*, vol. 21, no. 1, p. 18, Dec. 2020.

[149] V. Gunasegaran and V. Muralidharan, "Fault Diagnosis of Spur Gear System through Decision Tree Algorithm Using Vibration Signal," in *Materials Today: Proceedings*, 2019, vol. 22, pp. 3232–3239.

[150] M. Strączkiewicz and T. Barszcz, "Application of artificial neural network for damage detection in planetary gearbox of wind turbine," *Shock Vib.*, vol. 2016, 2016.

[151] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," *IEEE Signal Process. Mag.*, vol. 8, no. 4, pp. 14–38, 1991.

[152] H. Zheng, Z. Li, and X. Chen, "Gear fault diagnosis based on continuous wavelet transform," *Mech. Syst. Signal Process.*, vol. 16, no. 2–3, pp. 447–457, 2002.

[153] Y. Xu, C. Li, and T. Xie, "Intelligent Diagnosis of Subway Traction Motor Bearing Fault Based on Improved Stacked Denoising Autoencoder," *Shock Vib.*, vol. 2021, 2021.

[154] H. Liu, J. Zhou, Y. Zheng, W. Jiang, and Y. Zhang, "Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders," *ISA Trans.*, vol. 77, pp. 167–178, Jun. 2018.

[155] W. F. Wang, X. H. Qiu, C. Sen Chen, B. Lin, and H. M. Zhang, "Application Research on Long Short-Term Memory Network in Fault Diagnosis," *Proc. - Int. Conf. Mach. Learn. Cybern.*, vol. 2, pp. 360–365, Nov. 2018.

[156] X. Zhang, Y. Cong, Z. Yuan, T. Zhang, and X. Bai, "Early Fault Detection Method of Rolling Bearing Based on MCNN and GRU Network with an Attention Mechanism," *Shock Vib.*, vol. 2021, 2021.

[157] J. Grezmak, P. Wang, C. Sun, and R. X. Gao, "Explainable Convolutional Neural Network for Gearbox Fault Diagnosis," *Procedia CIRP*, vol. 80, pp. 476–481, Jan. 2019.

[158] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, 27-May-2015.

[159] C. Lu, Z. Wang, and B. Zhou, "Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification," *Adv. Eng. Informatics*, vol. 32, pp. 139–151, Apr. 2017.

[160] J. Jiao, M. Zhao, J. Lin, and J. Zhao, "A multivariate encoder information based convolutional neural network for intelligent fault diagnosis of planetary gearboxes," *Knowledge-Based Syst.*, vol. 160, pp. 237–250, Nov. 2018.

[161] G. Dalpiaz, G. Dalpiaz, A. Rivola, and R. Rubini, "Dynamic modelling of Gear system for condition monitoring and diagnostics," *Proc. Congr. Tech. DIAGNOSTICS*, p. 185192, 1996.

[162] M. D. Lutovac, D. V. Tošić, and B. L. Evans, *Filter design for signal processing using*

MATLAB and mathematica. Prentice-Hall: Englewood Cliffs, NJ, 2001.

[163] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2013, pp. 8609–8613.

[164] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998.

[165] C. C. Aggarwal, "Convolutional Neural Networks," in *Neural Networks and Deep Learning*, Cham: Springer International Publishing, 2018, pp. 315–352.

[166] Y. Yoo and J. G. Baek, "A novel image feature for the remaining useful lifetime prediction of bearings based on continuous wavelet transform and convolutional neural network," *Appl. Sci.*, vol. 8, no. 7, 2018.

[167] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017.

[168] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," 2014.

[169] C. D. Nguyen, A. Prosvirin, and J.-M. Kim, "Fault Identification of Multi-level Gear Defects Using Adaptive Noise Control and a Genetic Algorithm," in *Intelligent Human Computer Interaction*, 2021, pp. 325–335.