

FPGA BASED BEVEL GEARBOX FAULT DIAGNOSIS APPROACH

Nermen Salama, Rania Darwish, A.M. Bassiuny

Department of Mechatronics Engineering, Faculty of Engineering, Helwan

University,

Helwan, Egypt

E-mail: nermen_5@hotmail.com

Abstract— This paper addresses the issue of classifying local gear faults that depends on vibration signal measurements. Most of the early gears fault detection and diagnosis methods result insufficient results when dealing either with time or frequency domain characteristic. In order to overcome these obstacles, a Time-Frequency based approach implemented on Field Programming Gate Array (FPGA) is proposed. The presented approach combines Bagged Trees Classifier (BTC) with Complex Analytic Wavelet Transform (CAWT) analysis. Moreover, the presented approach benefits from the superior realization nature of FPGA. Extensive simulations and experiments have been conducted in order to demonstrate the efficiency of the proposed approach. Experiments are performed on bevel gearbox with both normal and one missing tooth. The obtained results clarify that the presented approach has a superior classification accuracy rate over the other comparative approaches.

Index Terms— Bevel gears, fault detection, complex analytic wavelet transform, vibration analysis, bagged trees algorithm, feature extraction, FPGA.

1. INTRODUCTION

Vibration measurements are regularly used in the rotating machine diagnostics field since changes in the condition of the machine are instantaneously reflected in the vibration[1]. However, changes in the machine condition (i.e. developing faults) are usually concealed by dominant components within the vibration signal, such as the

gear mesh frequencies, changes in operating conditions and noise. Sophisticated signal processing techniques are required to extract the characteristics or machine condition features that reflect the condition of the machine from the vibration signal.

Recently, Preventive Maintenance (PM) approach has attracted substantial attendance from many researchers from the field of vibration signal analysis. It assesses to maximize life cycle and cost optimization of rotary machinery components, also the early detection can get to maximize production capacity.

As all mechanical components are not run forever, failure in gearbox will directly out the machine from the service, which directly impacts on the production line. Sometimes it can stop the production line completely. Thus improving the reliability of gear could play a critical role in preventing unplanned failure.

Unexpected breakdown made the maintenance inefficient and costly for asset management. To overcome this situation, the prevented maintenance is used. Preventive maintenance involves the systematic inspection of equipment where potential problems are detected and corrected in order to prevent equipment failure. The PM is maintenance that is regularly performed on a piece of equipment to lessen the likelihood of it failing. PM is performed while the equipment is still working so that it does not break down unexpectedly[2]. The maintenance is scheduled based on a time or usage trigger. The frequency of preventative maintenance is most likely to be too high, which sacrificing reliability.

Time domain analysis based on statistics features such as the Crest Factor (CF), kurtosis, and Root Mean Square (RMS) can be used to capture changes in the machine condition due to an evidence of a fault. In [3], the authors combine the Fast Dynamic Time Warping (Fast-DTW) as well as the Correlated Kurtosis (CK). The work in [3] shows, a significant increase in its value in presence of gear tooth fault. The CK sensitive to the fault signature. In [4], the authors use condition indicators (RMS, peak) provide information that something has happened at variance with order analysis provides information about what has happened. In addition, it figures out that the condition indicator in fault beginning of the gearbox does not track the condition of the gearbox well during the test, e.g., the crest factor. Investigations on (FM0, FM4, NA4, NB4, and NA4) shows FM0 follows the trend of the crest factor. The trend of FM4, NA4, and NB4 are similar, which reflect how difficult to clarify the cause of the changes within the features which make the time domain features inappropriate for diagnostics in fluctuating operating conditions.

The work done in[3] offers a significant results out from measurements done on vibration signal to detect the local gear faults, the combination between the Fast-DTW as well as the CK techniques approach to this new time-domain fault detection method combines of the proposed approach investigations done to check the performance and applicability using analytical and dynamic simulation on planetary gear systems. The mathematical modeling, simulation and experimental

results presented in this work approach a useful developing automatic diagnostic algorithms to be applied to industrial systems that include both fixed axis as well as epicyclical gearboxes.

IGBA, et.al. [5] focused on the use of peak and RMS values of vibration signals for two different gearboxes. Three algorithms signal correlation, extreme vibration, and RMS intensity models, which developed and validated using CM data logged from operational gearboxes. Furthermore, data collected from 10 different gearboxes to test its detectability of common failure modes in the gearbox. The results showed that signal correlation with RMS values is good for early detecting progressive failures such as bearing pitting or shaft cracks. On the other hand, this was not executable for detecting gear tooth fracture. Unlike RMS, the peak values were better at detecting gear tooth fractures using both the correlation and extreme vibration model.

Ability to effectively represent and analyze data made by wavelets became more popular in mathematics and digital signal processing area. Wavelet basis functions are short non-zero oscillations in the time domain that are stretched (known as dilation) and translated. The wavelet transform is a time-scale distribution obtained by convolving classification ability and translated the vibration signal of interest. It must be remembered that the major concern is the description, representation, and reconstruction of functions. The time-scale distribution is well suited for non-stationary signal analysis, denoising and singularity detection [6].

The work done in[7] concerned with gear failure diagnosis based on wavelets, and image processing system for automatic fault detection and isolation. Various objects extracted from a set of product images the region of interest and some parameters calculated like area, major/minor axis length, orientation, diameter, convex area, Euler number. With a comparative study processed by time-domain waveforms, frequency spectrum, and Continuous Wavelet Transform (CWT) of the gearbox. Experimental analyses were done on gear and bearing in[8] using acoustic and vibration signals with near 100% accuracy with automatic nature in real-time monitoring practice. Filters applied on the geared system performs effectively in providing the information about multiple faulty teeth such as the angle between two damaged teeth using both kinds of signals. Filters applied on bearing, it's figured out that the severity of defect cannot be estimated adequately using vibration and acoustic signals. The proposed system of filters has diverse stages of signal processing such as denoising, time-frequency analysis, and extraction of Smooth Envelope Signal (SES) followed by a sturdy peak detection technique. Wavelet Packet Transform (WPT) has been exploited to the denoised signal.

An improvement did over classical Linear Discriminant Analysis (LDA) called , QR decomposition for Linear Discriminant Analysis (LDA/QR) introduced in[9]. It's a two-stage algorithm. LDA/QR has a highly efficient way to maximize the

separation between different classes, as a first step. Secondly, keeping low time/space complexity as low as possible while addressing the issue of within-class distance. LDA/QR algorithm made a scalable solution over both classical LDA and a combination of LDA and Principal Component Analysis (PCA) techniques. A comparative study was done in [10] between Euclidean Distance Technique (EDT) and the Fisher Discriminant Analysis (FDA) applied on a planetary gearbox test rig. With 19 feature parameters, its figure out that FDA has a curial reflection on classification accuracy.

In general, signal processing techniques such as time domain analysis, frequency analysis, and time-frequency analysis identified machine component faults. Fast Fourier Transform (FFT) converts time domain to frequency domain. FFT can be used to indicate the intensity of the frequencies in the spectrum. FFT cannot find the non-stationary transient information from the samples, which serves as the reason for focusing on the wavelet transform in this paper.

This paper focused on local faults in bevel gear especially one missing tooth. For simplicity, Different speeds and constant load are considered. Also, there are some factors can't be controlled perfectly, like temperature, humidity, the degree of wear and installation conditions. It must be noted that the learning process of BTC done in an offline condition. The figure (1) describes the workflow of the learning process.

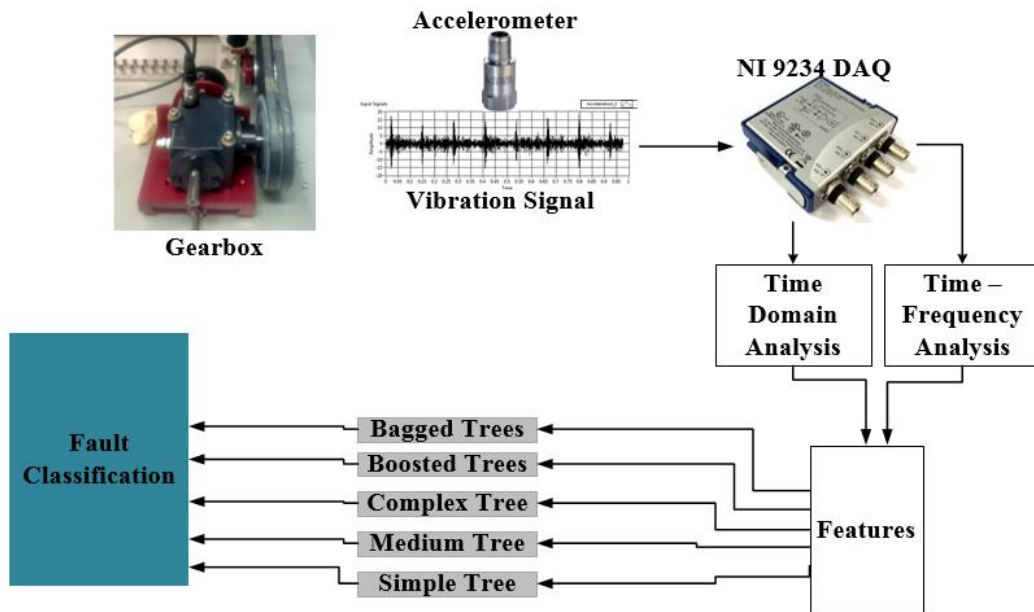


Figure 1 Offline Workflow Diagram

2. Analytic Wavelet Signal Analysis

Wavelet means “small wave”. Which mean a short duration finite energy. Hence, it should be time-limited. The wavelet ψ is a function of zero average having energy concentrated in time as seen in equation (1).

$$\int_{-\infty}^{\infty} \psi(t) dt \tag{1}$$

In order to extract more time- frequency information, a set of wavelets can be generated from a mother wavelet $\psi(t)$, which is limited to finite space. It exists only for a short duration and releases a finite amount of energy. Daughter wavelets $\psi_{a_1,a_2}(t)$ are generated by translation and dilation of wavelets, with parameters ‘ a_1 ’ and ‘ a_2 ’ correspondingly, equation (2).

$$\psi_{a_1,a_2}(t) = \frac{1}{\sqrt{a_1}} \psi\left(\frac{t - a_2}{a_1}\right) \tag{2}$$

The continuous wavelet coefficients, denoted by W , are obtained from the CWT, calculated from, equation (3)[11].

$$W(a_1, a_2) = \frac{1}{\sqrt{a_1}} \int_{-\infty}^{\infty} x(t) \psi^*\left(\frac{t - a_2}{a_1}\right) dt \tag{3}$$

Where the complex conjugate of the mother wavelet basis function and the investigated signal are denoted by ψ^* and $x(t)$, respectively. The proposed popular basis function for fault detection in rotary machines is Morlet is presented in figure (2), where the scaling and wavelet function is the low-pass and high-pass filter, respectively.

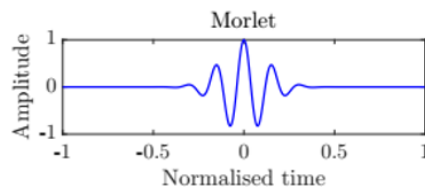


Figure 2 Morlet Wavelet Basis Function

The Analytic Wavelet Transform AWT is a distinguished condition of the CWT with the complex-valued Morlet wavelet, also it named Gabor wavelet. Equation (4) defines the complex-valued Morlet wavelet:

$$W_0(t) = \frac{1}{\sqrt[4]{\delta^2 \pi}} \exp\left(-\frac{t^2}{2\delta^2} + j\omega_0 t\right) \tag{4}$$

Where δ represents the standard deviation of the Gaussian envelope of the mother wavelet, and ω_0 represents the central frequency of the mother wavelet, which is $\frac{3}{2}\pi$. Therefore, the center frequency of the scaled wavelet is $\frac{\omega_0}{a}$ which effected on the center frequency of the scaled wavelet is inversely proportional to the scale a .

The AWT computes the inner products of the analyzed signal and a set of Complex Morlet Wavelets (CMW). This transform is called the AWT because the CMW are analytic, the power spectra of the Morlet wavelets are zero at negative frequencies. The resulting AWT coefficients are complex numbers. These coefficients measure the similarity between the analyzed signal and the CMW. The AWT is just one type

of complex continuous wavelet transform. It must be noted that the time and frequency resolutions of wavelets are adaptive, the AWT provides adaptive time and frequency resolutions. Conventional time-frequency analysis methods, such as the Short-Time Fourier Transform (STFT), only provide uniform time and frequency resolutions in the whole time-frequency domain.

3. Feature Extraction

The information about the health of the monitored Spectra-Quest Machinery Fault Simulator (MFS) is contained in the vibration signal. Statistical analysis techniques involve the extraction summarizing of features from the vibrational signature data could be used [12], and [13]:[15]. Ideally, these features are more stable and well behaved than the raw signature data itself. Features also provide a reduced dataset for the application of fault pattern recognition techniques. Before any feature can be calculated on the raw vibration data, the data must be conditioned or preprocessed. In this study conditioning based on the data acquisition unit. The features that used in this work were in time and time-frequency domain: Kurtosis, mean, R.M.S, peak, and CF. Figure (3) describes vibration-based gearbox condition analysis indicators flowchart.

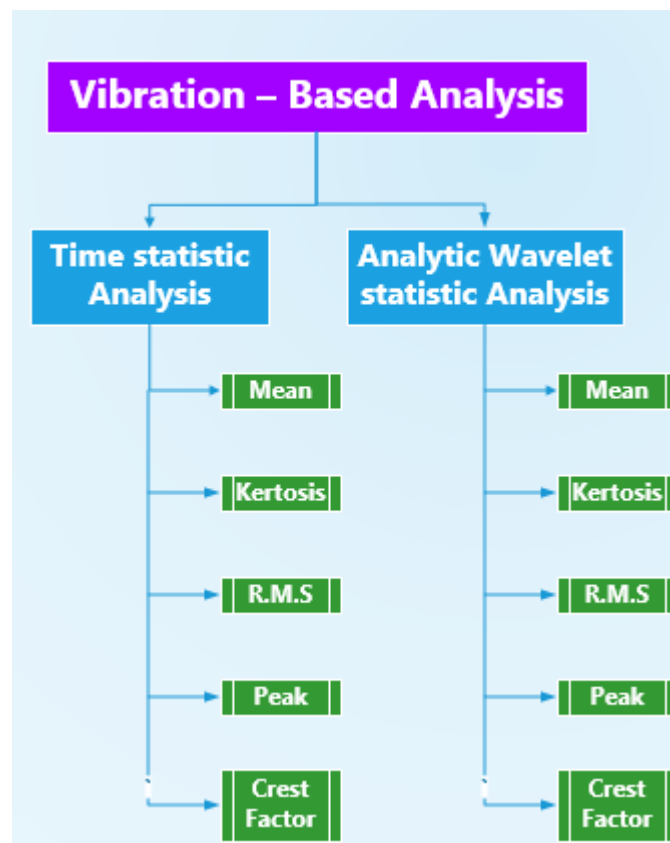


Figure 3 Vibration-Based Gearbox Condition Analysis Indicators Flowchart

3.1.1. Kurtosis

Kurtosis defined as the fourth moment of the distribution and measures the relative peakedness or flatness of a distribution as compared to a normal distribution. Kurtosis is given by equation (5)[12]:

$$K = \frac{\sum_{n=1}^N [y(n) - \mu]^4}{N * (\delta^2)^2} \quad (5)$$

Where $y(n)$ is the raw time series at point n , μ is the mean of the data, δ^2 is the variance of the data, and N is the total number of data points.

Kurtosis values can be summarized by equation (6).

$$Kurtosis = \begin{cases} \approx 3, & \text{normal – peak distribution} \\ < 3, & \text{flatter distribution than normal distribution} \\ > 3, & \text{sharper distribution than normal distribution} \end{cases} \quad (6)$$

3.1.2. Mean

The sample mean is an average value found in a sample. Mean is given by equation (7):

$$\bar{x} = \frac{\sum x_i}{n} \quad (7)$$

Where \bar{x} is the sample mean, x_i is all of the x-values, and n is the number of items in the sample.

3.1.3. Root Mean Square (RMS)

RMS describes the energy content of the signal and uses to evaluate the overall condition of the components. RMS is given by equation (8) therefore, it is not very sensitive to an incipient fault but used to track general fault progression[4].

$$X_{RMS} = \sqrt{\frac{1}{N} \sum_{i=0}^N (X_i^2)} \quad (8)$$

Where:

X_{RMS}	Root Mean Square value of dataset X .
X_i	i -th member of dataset X .
N	Number of points in dataset X .

3.1.4. Peak

The peak value is the maximum amplitude of the signals within a certain time interval. The peak can be calculated using equation (9).

$$X_{Peak} = \max(X_1, X_2, X_3, \dots, X_N) \quad (9)$$

3.1.5. Crest Factor

Crest Factor (CF) is the ratio of the single side peak value of the input signal to the RMS level[4]. CF can be calculated using equation (10).

$$CF = \frac{X_{Peak}}{X_{RMS}} \quad (10)$$

4. Feature selection:

Dimensionality reduction techniques transform the current feature space to a lower dimensional feature space by retaining most of the information content of the signal with respect to some criterion. The intrinsic dimension of the feature space can be significantly lower than the apparent or raw dimension of the features. This has a significant impact on the optimization of the models used to describe the feature space[16].

Linear Discriminant Analysis (LDA) aims to better discriminate patterns of different classes. For this distinguish characteristic, LDA is particularly suitable for solving problems of classification tasks. However, LDA pays no attention to the decorrelation of the data[17]. Lately, research community prefers LDA over Principal Component Analysis (PCA) because, LDA deals directly with discrimination between classes. On the other hand, PCA deals with the data in its entirety with poor attention to the underlying class structure[18]. As seen in figure (4) which adopted from [14], the LDA is in contrast with PCA as it is a supervised classifier. LDA identifies class membership, which leads to maximizing class separation.

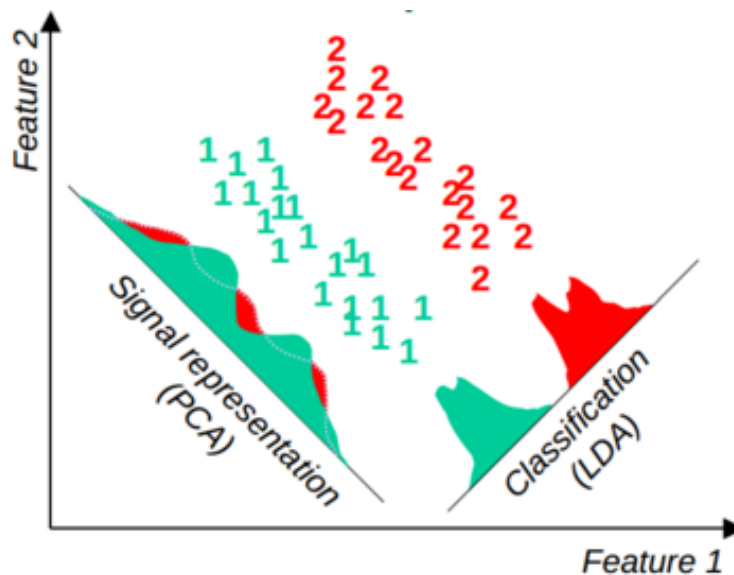


Figure 4 Illustrating The Different Concepts Between LDA and PCA[14]

Fisher-LDA considers maximizing the objective J using equation (11).

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}} \quad (11)$$

Where $J(\mathbf{w})$ is Fisher discriminant, \mathbf{w} represents the linear transformation that maps the original t-dimensional space onto a f-dimensional feature subspace. \mathbf{S}_B is the between classes scatter matrix, and \mathbf{S}_W is the within classes scatter matrix which can be defined as equations (12), and (13) where $\bar{\mathbf{x}}$ is the overall mean of the data-classes and μ_c is the mean of class. Because the scatter matrices are proportional to covariance matrices, it can define objective using

$$\mathbf{S}_B = \sum_c (\mu_c - \bar{\mathbf{x}})(\mu_c - \bar{\mathbf{x}})^T \quad (12)$$

$$\mathbf{S}_W = \sum_c \sum_{i \in c} (\mu_c - \bar{\mathbf{x}})(\mu_c - \bar{\mathbf{x}})^T \quad (13)$$

covariance matrices.

In addition, projected data can be defined as equation (14).

$$\mathbf{y} = \mathbf{w}^T \mathbf{x} \quad (14)$$

To solve equation (15), a good solution is one where the class-means are well separated, measured relative to the sum of the variances of the data assigned to a particular class. This is precisely the objective because it implies that the gap between the classes expected to be big. Which adopt \mathbf{w} such that the denominator is $\mathbf{w}^T \mathbf{S}_W \mathbf{w} = \mathbf{1}$, since \mathbf{w} is a scalar itself.

$$\mathbf{w} = \mathbf{S}_W^{-1}(\mu_c - \bar{\mathbf{x}}) \quad (15)$$

5. Machine learning

Decision Tree

In supervised machine learning, ensemble is popular because of the ability to accurate prediction. An ensemble method uses more than one classifier overall better accuracy. Classical ensemble bagging and boosting, have good predictive cap method was proposed by work in [19]. Study done in [20] shows how Bagging Tree is able to identify two different diseases like Myocardial Infarction (MI), and Arrhythmia (AR) from normal patients with the cross-validation accuracy reaches 99.7%, and sensitivity near 99.4%, specificity of above 99.5%, with a precision of 99.32%, and F1 gets a score of 99.36% from a single lead electrocardiography data. The proposed

algorithm can be realized in wearable form to fetch a real-time patient monitoring. The quality of sleep in the work[21] has a major impact on human health and his life quality.

A classification based on bagged trees for sleep stages collected with a Continuous Wave (CW) Doppler radar investigated in this work. Experiments were done on people who asked to sleep all night with polysomnography. Nine labeled features extracted from the radar signals.

Four kinds of tree classifiers were compared as the machine learning algorithms to classify sleeplessness, Rapid Eye Movement (REM) sleep, light sleep and deep sleep. Testing and validating done with 10-fold cross-validation procedure was used to address the classifier performance.

As a result, the bagged trees classifier has the best classification accuracy rate over the four classifiers. Which lead to improvement in accuracy rate can be reached to 78.6%.

In bagging algorithm form original training data set N different samples called bootstrap samples [19] $X_1, X_2 \dots X_n$ are generated. A classifier C_n is built against sample S_i . From classifier $C_1, C_2 \dots C_n$, final classifier C_p is built whose output is the class predicted most classifiers. The bagging process is shown in figure (5) which adopted from[19].

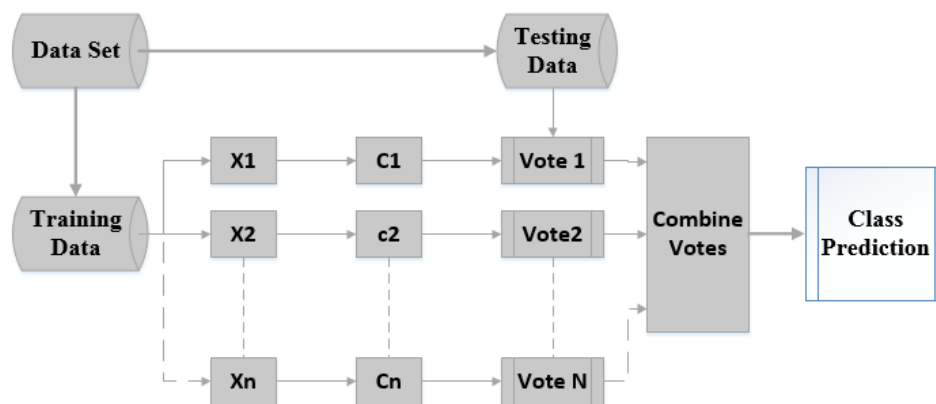


Figure 5 The Bagging Tree Flow Chart

Bagging trees are based on the bootstrap sampling. From the original dataset, we choose the S samples randomly as a training set which is the base learner. With N times repeating, we get a concord classifiers. The result is got from lump vote equation (14), which is introduced as follow:

$$C(x) = \underset{j}{\operatorname{argmax}} \sum_{i=1}^n C_i^j(x) \tag{14}$$

Where $C_i^j(x)$ is the result for n classifier. Table (1) summarizes the pseudocode for the Bagging Trees algorithm.

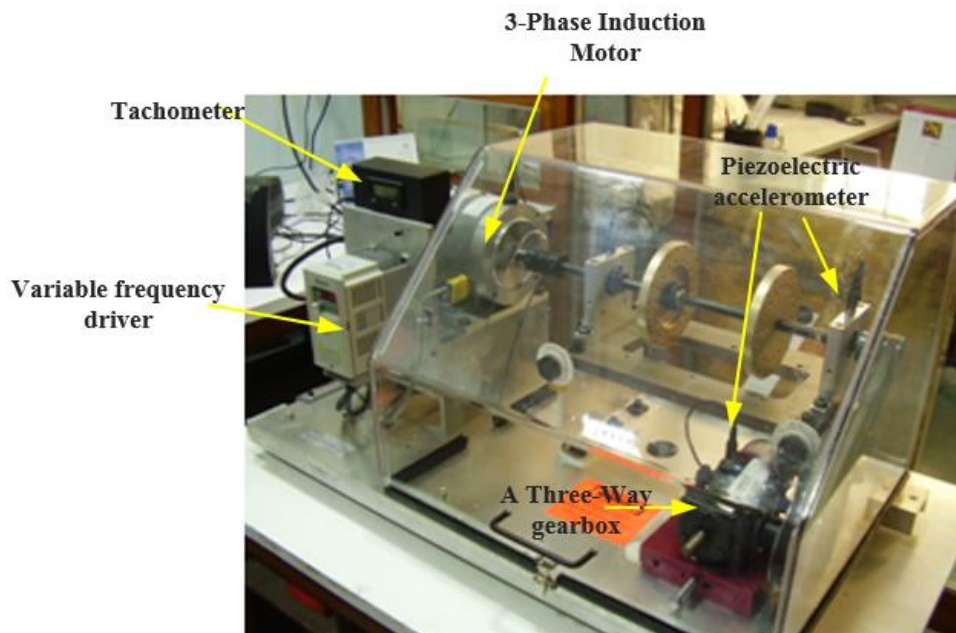
Table 1 The Bagging Trees Algorithm Pseudocode

-
- 1: **Procedure:** Train.
 - 2: **While** Training time < N .
 - 3: $D \leftarrow$ Training Data set.
 - 4: $Y \leftarrow$ Output for input X .
 - 5: $\sigma \leftarrow$ Base Learning Algorithm.
 - 6: $N \leftarrow$ Training time.
 - 7: $D = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_i, Y_i)\}$
 - 8: **For** $i = 1$ **to** n .
 - 9: $C = \sigma(D)$
 - 10: **Result:**
 - 11: $C(x) = \underset{j}{\operatorname{argmax}} \sum_{i=1}^n C_i^j(x)$

6. Experimental Setup

The experimental setup [22] is shown in figure (6) consists of a 3-phase, 1 HP variable speed induction motor, controlled manually by variable frequency driver with multi-featured front panel programmable controller. The motor speed can be varied from 0 to 3600 rpm. A three-way straight cut bevel gearbox with a 1.5:1 ratio is shown in figure (7) (a) and (b).

The gearbox specifications are presented in table (2). It must be hinted that the gears were made from forged steel and the bearing used is (NSK 6202).

**Figure 6 Experimental Setup**

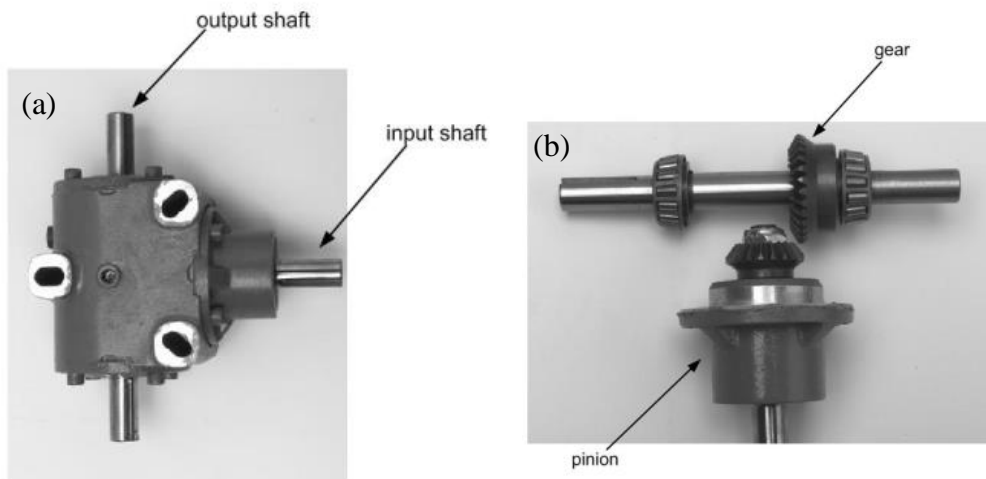


Figure 7 (a) A Three-Way Gearbox And (b) Internal Construction

Table 2 Gearbox Specification

	Pinion	Gear
<i>Pitch angle</i>	33° 41'	56° 19'
<i>Pressure angle</i>	20°	20°
<i>Pitch diameter</i>	2.8575	4.28625
<i>Number of teeth</i>	18	27

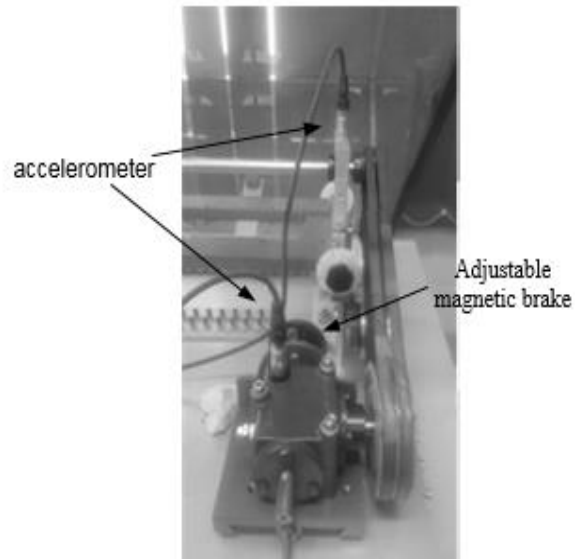


Figure 8 Accelerometer locations

Data collected from two locations as seen in figure (8), Bearing and gearbox housing. An adjustable magnetic brake working as a torque load ranging from (0.0565 to 1.1298 N.m) connected to the gearbox, which gets the power from two double grooves V-belt attached to a 25.4 mm shaft diameter with a torsional stiffness spring coupling. The vibration signal captured from the test rig by ICP (603C01) piezoelectric accelerometer with 104 mV/g actual sensitivity. Twice, firstly the data logged using (NI 9234 card) which connected to the laptop for online monitoring to test and handle problems faced in the begin of the experimental work algorithms. Secondly, it made a stand-alone system using Compact Reconfigurable Input Output (cRIO-9074) controllers which based on the LabVIEW RIO architecture, featuring a processor, reconfigurable FPGA, and I/O interface. Including programming environments LabVIEW[23]. Figure (9) shows faulty gears

with 100% one removed tooth. Five input speeds (15rps, 20rps, 25rps, 30rps, and 35rps) with constant load condition.

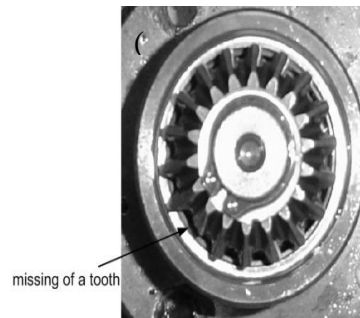


Figure 9 One Missing Tooth Gear.

7. Results:

As for one missing tooth fault seen in figure (10), it has been figured out that data collected from Gearbox casing was more cleared than data collected from bearing housing. Also, it can be illustrated by the same figure that the power of impacts drops and noise increases as a natural result of power loss during transmitting with the pulley-belt mechanism. As a result, this paper focused on studying the vibration signals collected from the gearbox casing position. Vibration signals acquired by the FPGA, then processing by power spectrum using the FPGA and transmitted to the real-time embedded processor to get Analytic Wavelet, are shown in figures (11) (a) and (b), (12) (a) and (b), and (13) (a) and (b).

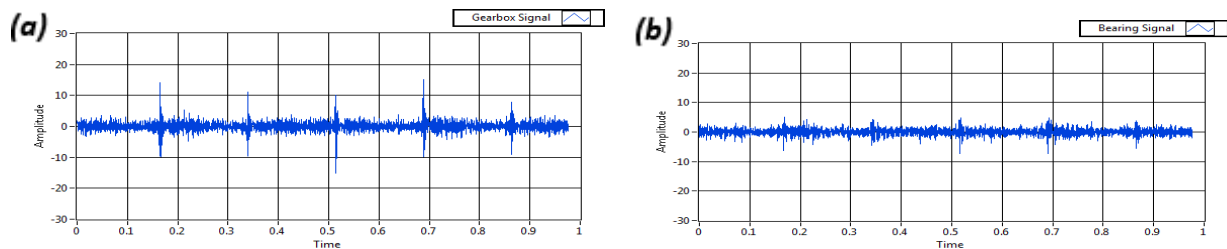


Figure 10 One Missing Tooth Signal From (a) Gearbox Casing and (b) Bearing Housing

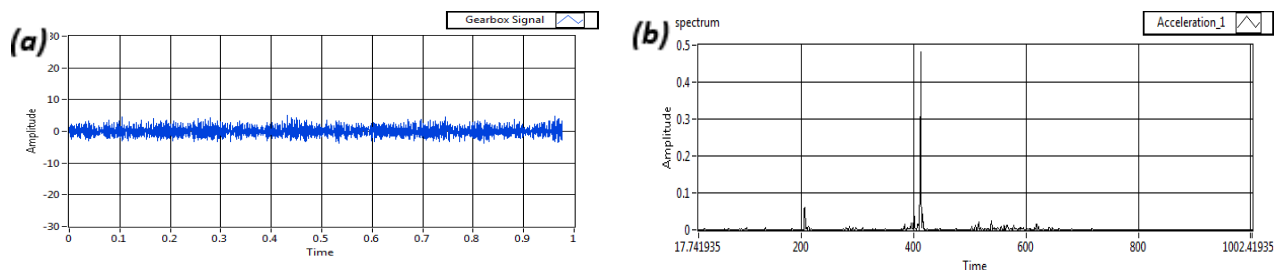


Figure 11 Normal Condition (a) Time Domain Vibration Signal and (b) Power Spectrum

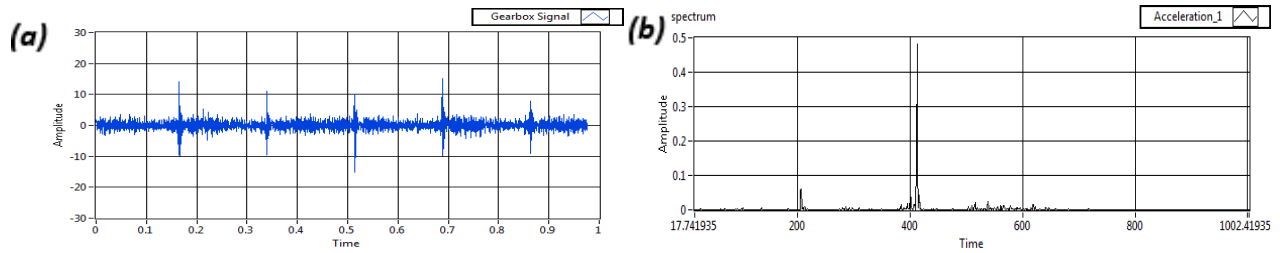


Figure 11 One Missing Tooth (a) Time Domain Vibration Signal and (b) Power Spectrum

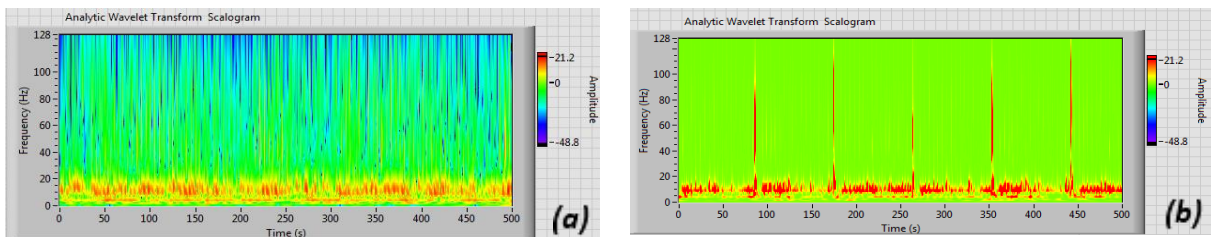


Figure 13 Both (a) Normal Condition and (b) One Missing Tooth Gearbox Analytic Wavelet

For the sack of simplicity, in this work some parameters are assumed to be fixed, like splitting criteria, a number of splits, with no surrogate decision splits for simple, medium, and complex trees. Also, some fixed parameters did on Boosted, and Bagged Trees like, the maximum number of splits, learning rate, and number of learners. Table (3) recognize this parameter. The recognition rate for Trees can be clear in figure (13).

Table 3 Fixed Parameters Taken in The Experiments Done by Decision Trees

Decision Tree	Split criterion	Max number of splits	Learning Rate	Number of Learner	Learning Consumed Time (Sec)
Simple Tree	Gini's diversity index	4	-	-	0.01
Medium Tree		20	-	-	0.00
Complex Tree		100	-	-	0.00
Boosted Trees	AdaBoost	20	0.1	30	0.02
Bagged Trees	Bag				0.08

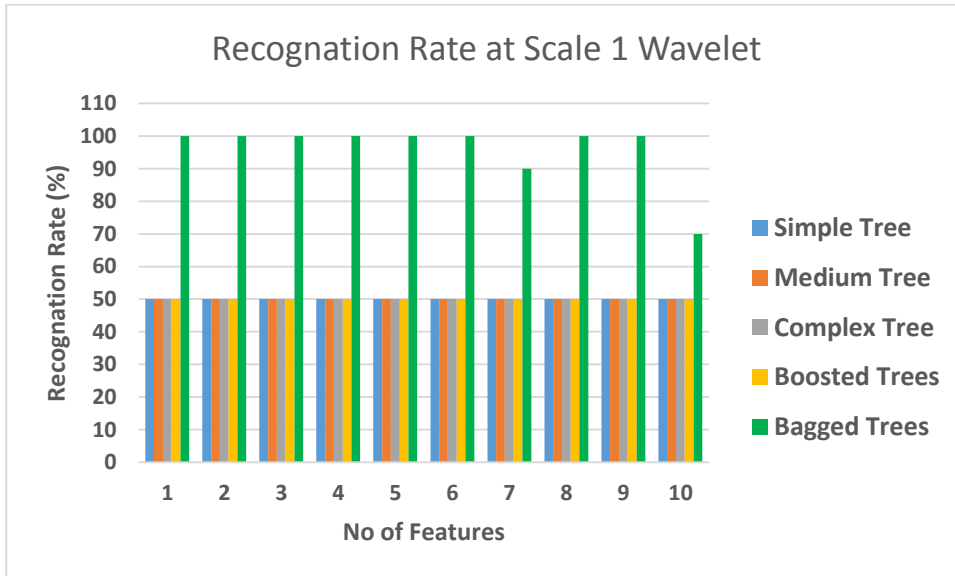


Figure 13 Recognition Rate of Bagged Decision Tree at Scale 1 Wavelet

The selection of Compact RIO target FPGA consider the required resources for implementing the approached algorithm, the result provides true parallel processing with the performance and reliability benefits of FPGA implementing circuitry. Because there is no operating system on the FPGA chip, the code is implemented in a way that ensures maximum performance and reliability.

The architecture software application is designed under a hierarchical form figure (14). LabVIEW environment allows code development for FPGA. Also, the real-time processor communicates internally with the FPGA using a high-speed bus circuit.

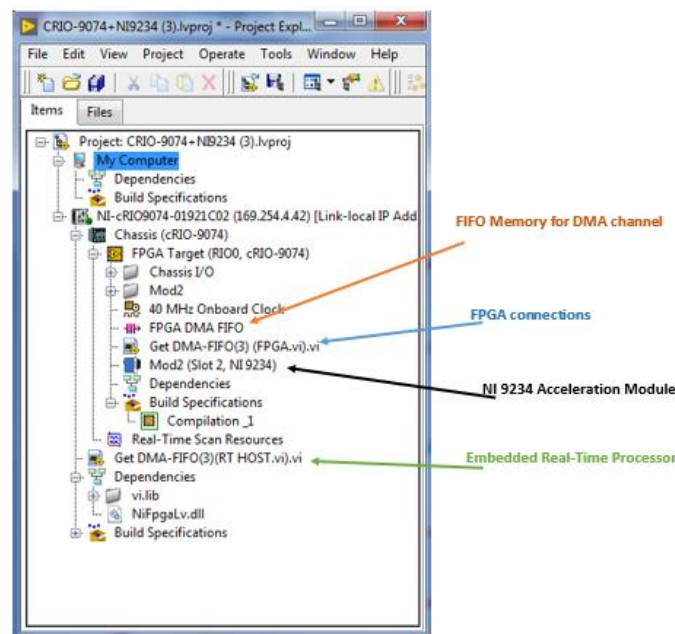


Figure 14 Organizing Of The LabVIEW Application.

The acquired information is analyzed by the proposed accelerometer sensor that is implemented in proprietary Xilinx Spartan 3 2M NI FPGA platform running at 40 MHz Table (4) summarizes the resource usage of the FPGA.

Table 4 Resources Usage of The FPGA

<i>Resource utilization</i>	<i>Xilinx Spartan</i>
<i>Slices</i>	2457/20480 (12%)
<i>Flip-flops</i>	820/40960 (2%)
<i>4-input LUTs</i>	4502/40960 (11%)
<i>Maximum operation frequency</i>	53.012 MHz

8. Validate the Purposed System

The proposed detection methodology used in this work was validated using experimental gearbox data collected from industrial straight cut bevel gearbox as seen in figures (15) (a) and (b). Two industrial straight cut bevel gearboxes are used to validate the proposed technique, with both normal and one missing tooth conditions. The gearboxes driver and driven had 11, and 38 teeth respectively.

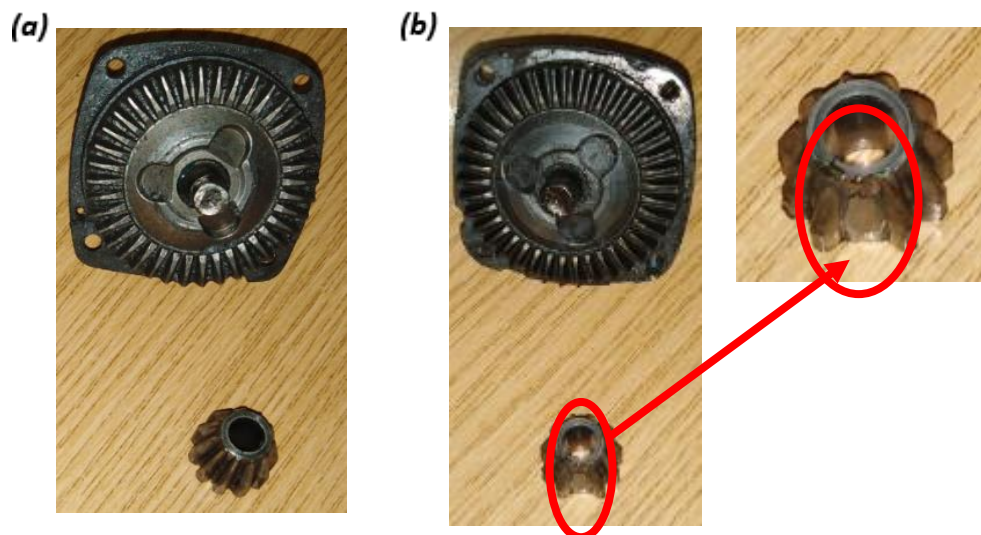


Figure 15 Driver and Driven Gears in Both (a) Normal Condition and with (b) One Missing Tooth Condition

An artificial tooth reshaping was created by grinding the tooth completely of the driver gear to simulate the missing tooth scenario as shown in figures (16) (a) and (b) .

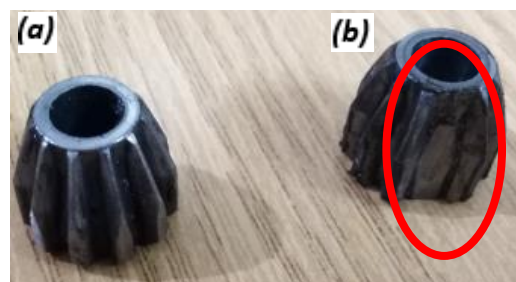


Figure 16 Driver Gear in (a) Normal Condition and with (b) One Missing Tooth

9. Conclusion

This research work introduced local bevel gearbox faults classifier that depends on vibration signal measurements through an FPGA-based implementation that provides an efficiently real-time online portable application to condition monitoring community. The conclusions of this work can be summarized as follows:

1. The Increase in the frequency scale of the Analytic Wavelet Transform (AWT) tends to decrease the center frequencies of the wavelets which in turn reflected on the computation time of AWT. In another way, the increasing scale of AWT leads to increase in computational time.
2. Changing the frequency level of the AWT has no effect on the performance of the classifier.
3. Bagged decision trees classifier showed its superiority over the backpropagation neural network in both training time and a number of features used.
4. The strength of the signal remains at the same level even if the number of broken teeth increases however, it increases the number of peaks.
5. Contrary to expectations from literature review results showed that LDA algorithm is much better than the PCA algorithm in features selection.
6. The strength of the signal increases as a result of increasing the speed of gearbox.
7. AWT gives a great indication with respect to nonstationary operating conditions caused by tooth missing.
8. Increasing speed of the gearbox affect directly on the signal captured in the terms of amplitude, noise rates, and a number of cycles per time.
9. The use of the FPGA appears to be the best in terms of computational time compared to traditional multi-core Processor as expected.
10. With the help of a LabVIEW program and implementation on NI hardware, it can be easy to develop advanced techniques for fault detection embedded system.

10. Acknowledgments

The authors gratefully acknowledge use of the facilities provided through DANIDA grant and is available at HU-FLSmith Cement Technology lab.

Acknowledgment also to Mr. Sherif Ismail, Country Sales Manager - NI Egypt, who provided expertise that greatly assisted the research.

11. References

- [1] R. B. Randall, *Vibration-based Condition Monitoring*. Chichester, UK: John Wiley & Sons, Ltd, 2011.
- [2] Fixsoftware, “Preventative Maintenance (PM).” [Online]. Available: <https://www.fixsoftware.com/maintenance-strategies/preventative-maintenance/>. [Accessed: 10-Feb-2017].
- [3] L. Hong and J. S. Dhupia, “A Time Domain Approach to Diagnose Gearbox Fault Based on Measured Vibration Signals,” *J. Sound Vib.*, vol. 333, no. 7, pp. 2164–2180, 2014.
- [4] P. Večeř, M. Kreidl, and R. Šmíd, “Condition Indicators for Gearbox Condition Monitoring Systems,” *Acta Polytech.*, vol. 45, no. 6, pp. 35–43, 2005.
- [5] J. Igba, K. Alemzadeh, C. Durugbo, and E. T. Eiriksson, “Analysing RMS and peak values of vibration signals for condition monitoring of wind turbine gearboxes,” *Renew. Energy*, vol. 91, pp. 90–106, 2016.
- [6] J. Rafiee and P. W. Tse, “Use of Autocorrelation of Wavelet Coefficients for Fault Diagnosis,” *Mech. Syst. Signal Process.*, vol. 23, no. 5, pp. 1554–1572, Jul. 2009.
- [7] X. Zhao, “A Method of Gear Fault Detection Based on Wavelet Transform,” in *2009 International Conference on Business Intelligence and Financial Engineering*, 2009, no. 6, pp. 648–651.
- [8] D. P. Jena and S. N. Panigrahi, “Automatic Gear and Bearing Fault Localization Using Vibration and Acoustic Signals,” *Appl. Acoust.*, vol. 98, pp. 20–33, 2015.
- [9] J. Ye, S. Member, Q. Li, and S. Member, “A Two-Stage Linear Discriminant Analysis,” vol. 27, no. 6, 2005.
- [10] H. Li, J. Zhao, R. Yang, J. Zhao, and H. Teng, “Research on planetary Gearboxes Feature Selection and Fault Diagnosis Based on EDT and FDA,” *Proc. 2014 Progn. Syst. Heal. Manag. Conf. PHM 2014*, pp. 178–181, 2014.
- [11] NI, “Analytic Wavelet Transform,” 2010. [Online]. Available: http://zone.ni.com/reference/en-XX/help/371419D01/lvasptconcepts/wa_awt/. [Accessed: 30-Apr-2018].
- [12] M. Lebold, K. McClintic, R. Campbell, C. Byington, and K. Maynard, “Review of Vibration Analysis Methods for Gearbox Diagnostics and Prognostics,” *54th Meet. Soc. Mach. Fail. Prev. Technol.*, pp. 623–634, 2000.
- [13] J. Zhu, T. Nostrand, C. Spiegel, and B. Morton, “Survey of Condition Indicators for Condition Monitoring Systems,” *Annu. Conf. Progn. Heal. Manag. Soc.*, vol. 5, pp. 1–13, 2014.
- [14] R. Li and M. Frogley, “On-Line Fault Detection in Wind Turbine Transmission System using Adaptive Filter and Robust Statistical Features,” *Int. J. Progn. Heal. Manag.*, pp. 1–9, 2013.
- [15] G. Liao, T. Shi, W. Li, and T. Huang, “Feature selection and classification of gear faults using SOM,” *Lect. Notes Comput. Sci.*, vol. 3498, no. III, pp. 556–560, 2005.

- [16] C. M. Bishop and Pattern, *Pattern Recognition and Machine Learning*. Springer Science+ Business Media, LLC, 2006.
- [17] Y. Liang, C. Li, W. Gong, and Y. Pan, “Uncorrelated Linear Discriminant Analysis Based on Weighted Pairwise Fisher Criterion,” *Pattern Recognit.*, vol. 40, no. 12, pp. 3606–3615, Dec. 2007.
- [18] A. C. Kak, “PCA versus LDA,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 228–233, 2001.
- [19] S. A. A. Shah, W. Aziz, M. Arif, and M. S. A. Nadeem, “Decision Trees Based Classification of Cardiotocograms Using Bagging Approach,” *Proc. - 2015 13th Int. Conf. Front. Inf. Technol. FIT 2015*, pp. 12–17, 2016.
- [20] S. Khatun and B. I. Morshed, “Detection of Myocardial Infarction and Arrhythmia from Single-lead ECG Data using Bagging Trees Classifier,” in *2017 IEEE International Conference on Electro Information Technology (EIT)*, 2017, pp. 520–524.
- [21] L. Zhang, J. Xiong, H. Zhao, H. Hong, X. Zhu, and C. Li, “Sleep Stages Classification by CW Doppler Radar using Bagged Trees Algorithm,” *2017 IEEE Radar Conf. RadarConf 2017*, pp. 0788–0791, 2017.
- [22] I. SpectraQuest, *Machinery Fault Simulator*, vol. Package. 2011, pp. 4–8.
- [23] National Instruments, “NI cRIO-9074XT User Manual and Specifications,” 2016.