

Fault Diagnostics of a Gearbox with Acoustic Signals Using Wavelets and Decision Tree

Yasir Khan^{1*}, Siju Abraham¹, V. Sugumaran¹ and M. Amarnath²

¹School of Mechanical and Building Sciences (SMBS), VIT University, Chennai Campus, Chennai – 600127, Tamil Nadu, India; kyasir27@gmail.com, siju.abby@gmail.com, v_sugu@yahoo.com

²Indian Institute of Information Technology Design and Manufacturing Jabalpur, Jabalpur - 482005, Madhya Pradesh, India; amarnath@iiitdm.in

Abstract

Objectives: This study aims at devising a methodology for accurately predicting the different fault conditions of gears in a gearbox using acoustic signals. **Statistical Analysis:** The acoustic signals are captured for several artificially created fault conditions of different magnitude and the wavelet features are extricated from captured acoustic signals. Subsequently, prominent features are selected by utilizing J48 Decision tree which discerns the most dominant traits among the allocated data obtained from wavelet transform of the acoustic signals followed by Random Forest for the classification of features. **Findings:** Out of a total of eleven features extracted, six were selected through Decision Tree and Random forest was used for feature classification of acoustic signals using wavelet features. Several iterations were conducted on the wavelet features by varying different parameters and the maximum percentage accuracy was found to be 99.76%. The instances of misclassification of features were minimal in Random Forest and it proved to be an efficient and precise classifier. Hence, Random Forest proved to be an easy to use, fast and accurate classifier that could classify various kinds of wavelet features efficiently. **Applications:** The methodology can be used to provide accurate real time results about the condition of gear teeth.

Keywords: Acoustic Signals, Decision Tree, Fault Diagnostics, Gear Box, Wavelets

1. Introduction

The gearbox is the most vital component in the transmission/drive train of a vehicle or any multi-speed machine. The gearbox contains multiple gears which are attached to output, input and secondary shafts. These gears have multiple teeth on their periphery which remain in contact with each other while transmitting power. Hence, due to heavy frictional force and heavy loading conditions they tend to wear out periodically with time. The wear if not detected beforehand, can possibly lead to major failures in the machine and even risk human life in case of failure in vehicle gearbox. Therefore, it becomes all the more important to devise a methodology, that is, a real time conditioning monitoring system, which predicts the fault occurrence in a gear teeth precisely so that possible accidents can be avoided. As the gears are

enclosed in metal casing, the only possible way to extract information is through vibration and sound signals. This paper addresses the pressing need for real time condition monitoring fault diagnostics of a gear box via acoustics signals extracted from the gears.

Both vibration and acoustic signals can be used for study; however, acoustic signals have been used due to financial restrictions, an expensive high frequency response accelerometer would be required if study is carried out using vibration signals. On the other hand, simple microphone sensors work well for extracting acoustic signals. The acoustic signals thus obtained are used for further operation that provide quite accurate classification of gear tooth faults which help us in predicting the future faults in the gear box.

The fault diagnosis most commonly comprises of three major steps- feature extraction, feature selection and

* Author for correspondence

feature classification. The feature extraction techniques are categorized as- statistical features¹, histogram features² and wavelet feature³. The wavelet feature has been used in the following study. Many kinds of feature extraction techniques are available, of which Fast Fourier Transform (FFT) is one of the most commonly used. However it has a major drawback, i.e., it cannot extract enough features without sufficient samples. Also, it is not much helpful in extracting non stationary signals like from gear box⁴.

Another technique, Short Time Fourier Transform (STFT) which was initially brought to rectify the problems faced by FFT has its own downside being the particular size of time window chosen at early stages remains constant for all frequencies⁵. Hence, if the window is too short or not of required size, then each FFT result might detect not only the desired high frequency component of interest, but also a greater bandwidth of adjacent frequencies. The Continuous Wavelet Transform (CWT) has been used by Michael et al in his work on "Rapid computation of CWT by oblique projections"⁶. Although they concluded that the method is adaptable enough to estimate several wavelet shapes and can achieve a randomly fine sampling of the scale axis and is good for fault diagnostics, it takes immense computational time. Another technique, Discrete Wavelet Transform (DWT) had been used by Lorand et al in his work on "Discrete wavelet transform based rotor faults detection method for induction machines", which concluded that DWT can be used successfully for the rotor fault detection of wound rotor induction machines and also the application can also be extended to other mechanical components like gears and bearings⁷.

A study by Krishnakumari et al using DWT and Zhao-Atlas-Marks distribution on spur gear fault diagnostics concluded that Daubechies wavelet was producing the best features set in the research. The paper further stated that DWT has optimal frequency accuracy in discerning the defective gear at low frequency band width and DWT is better than STFT in discerning the defective gear from the normal gears. For feature selection, prominent techniques used are Decision Tree⁸, Genetic Algorithm⁹, and Principal Component Analysis¹⁰ etc. The main disadvantages of Genetic Algorithm are that it has no guarantee of finding global maxima. Also, it takes time for convergence. The Principal Component Analysis takes assumptions like the dimensionality of the given data can be effectively minimized through linear transformation and that the majority of information is held in directions where input data variance is highest.

These conditions are rarely met hence the method is not advisable. The paper contains the usage of J48 tree because it is compact, fast, and easy to use, and most importantly has the best selection accuracy. Finally, for feature classification, leading classifiers used are Proximity Support Vector Machine (PSVM)¹¹, Support Vector Machine (SVM)¹², Fuzzy¹³, Logistic Regression¹⁴, Naive Bayes¹⁵, Random Forest¹⁶ etc. The SVM has a major disadvantage of having a high algorithmic complexity and extensive memory requirements in large scale tasks¹⁷. Furthermore, it is a 2-class classifier and to overcome such limitations a combination of three classifiers needs to be used¹⁸. On the other hand, Naive Bayes has a strong feature independence assumption; furthermore it requires big data sets for accurate classification of features. Also, individual fuzzy if-then rules for each class are not always enough for real-world pattern classification problems²¹. Logical Regression requires much more data to achieve stable, meaningful results. To overcome the above disadvantages, this paper uses DWT energy based features with random forest algorithm. To select only contributing features, J48 decision tree algorithm was used.

2. Experimental Setup

The apparatus of the experiment contains a 5 horsepower helical gearbox that has two stages of operation. The helical gearbox is run by another 5.5 horsepower 3-phase induction motor at 1200 rpm. It produces a mechanical output which is used to drive a D.C generator. The DC motor produces 2kW power while working as a generator that dissipates to a resistor bank connected to it. This electrical setup is made in order to avoid the torsional vibrations which may occur if traditional dynamometers are connected. All three devices are installed on stiffened I-beams that are further attached to a concrete slab. A piezo-electric accelerometer with model no. B&K 4332 is fitted along the stud to calculate the signals from vertical vibration that are produced on the bearing housing of the 16 teeth pinion. The frequencies of gears to be meshed are measured at 320 Hz and its multiples. Different data sets were collected when the helical gear train was working at Good, 20%, 40%, 60%, 80%, 100% and 150% fault conditions. A total of 60 data sets were gathered for every operating condition. The signals were shortened to 3 kHz by applying a low pass filter and sampled at 8 kHz. The output from the B&K 4332 was conditioned using B&K 2626 charge amplifier as shown in Figure 1. As the

placement of microphone was very important in this experiment, different positions and directions are tried out and found to produce good reception when kept at a gap of 55mm from the pinion.

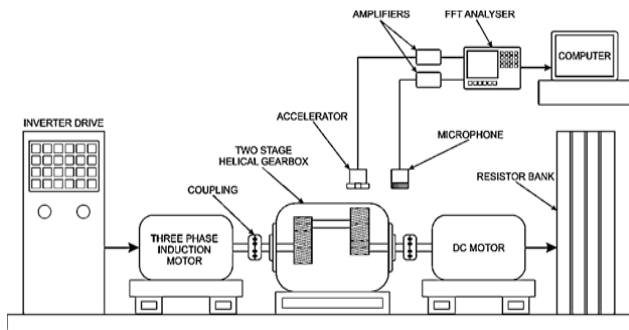


Figure 1. Experimental setup of two stage helical gearbox.

The pinion is connected to a DC motor to come up with two kilowatt power. Therefore, truth load on the turbine shell is around two. 6 HP that is barely fifty two of its rated power of five HP within the case of typical dynamometers, supplementary torsional vibrations could occur attributable to torsion fluctuations. Electrical device bank and DC motor are accustomed avoid such state of affairs to limit the recoil within the system to simply the gears, tire couplings are connected between the electrical machines and kit box.

The repair of another rigging box is required following a year. Seeded blame trials are basic to study blame discovery methodology. Along these lines, neighborhood blames in an apparatus box are sorted into: 1. broke tooth, 2. loss of a piece of tooth because of breakage of tooth at root or at a point on working tip, and 3. Surface wear spalling. There are different techniques to reenact blames in riggings like granulating, quickened test condition and EDM. Among them the most effortless strategy is called halfway tooth expulsion which reproduces the fractional tooth soften and is standard up a few modern applications.

3. Feature Extraction

The DWTs are a type of wavelet transforms for which the wavelets are discontinuously or discretely placed. A major advantage of DWTs over another kind of transform is the usage of temporal resolution, i.e., it can simultaneously capture both location and frequency.

The wavelet decomposition was performed using

DWT on vibration signals. The decomposition resulted in several trends and details. The trends were further decomposed into consecutive level trend and details. The trends of previous levels were subsequently decomposed and several levels of details were obtained. Here, the length of the signal is 2048 (2^{11}) and thus, 11 levels of decomposition are possible overall. At each level, the detail co-efficient were used to compute energy content using the following formula.

$$V_i = \sum_{i=1}^n x_i^2$$

Where x_i = details coefficients.

n = number of details coefficients.

The features were defined as the energy content at subsequent levels. The feature vector is defined as:

$$V = (v_1, v_2, v_3, \dots, v_m)$$

When m - (number such that length of signal) = 2^m

where v_1, v_2, v_3, \dots are energy content at given level.

Various families of wavelets are taken into account here. They are as follows:

- Haar wavelet
- Discrete Meyer wavelet
- Daubechies wavelet – Db1, db2, db3, db4, db5, db6, db7, db8, db9, db10
- Biorthogonal wavelet – bior1.1, bior 1.3, bior 1.5, bior 2.2, bior 2.4, bior 2.6, bior 2.8, bior 3.1, bior 3.3, bior 3.5, bior 3.7, bior 3.9, bior 4.4, bior 5.5, bior 6.8
- Reversed Biorthogonal wavelet – rbio1.1, rbio 1.3, rbio 1.5, rbio 2.2, rbio 2.4, rbio 2.6, rbio 2.8, rbio 3.1, rbio 3.3, rbio 3.5, rbio 3.7, rbio 3.9, rbio 4.4, rbio 5.5, rbio 6.8
- Coiflet – coif 1, coif 2, coif 3, coif 4, coif 5
- Symlets – sym 2, sym 3, sym 4, sym 5, sym 6, sym 7, sym 8

4. Wavelet Selection

The time domain signals were obtained using 54 distinct discrete wavelet transforms from the 7 wavelet families mentioned above. Consecutively, the extracted feature from each of the wavelet transform is put into J48 algorithm to find the maximum classification accuracy. The features extracted using Daubechies 5 provided the best classification accuracy by using J48 Decision Tree. It had been chosen for later operations. The Daubechies wavelets are orthogonal wavelets set apart by a maximal number of vanishing minutes for a specific support furthermore set up a discrete

wavelet change. With each wavelet sort, a scaling capacity is available which creates an orthogonal multi-determination examination. The Daubechies wavelet, represented as 'db n' is a family of orthogonal wavelets that have the maximum number of vanishing points(n) for any given support width of $2n-1$. The result is taken from the solution whose scaling filter is generating the highest phase from the $2n-1$ possible solutions for a point along with its orthogonality conditions. Therefore, the Daubechies 5 (db5) was selected among many DWT since it produced the best percentage accuracy when used with C4.5 algorithm. Figures 2-9 give the classification accuracy of various wavelet families individually.

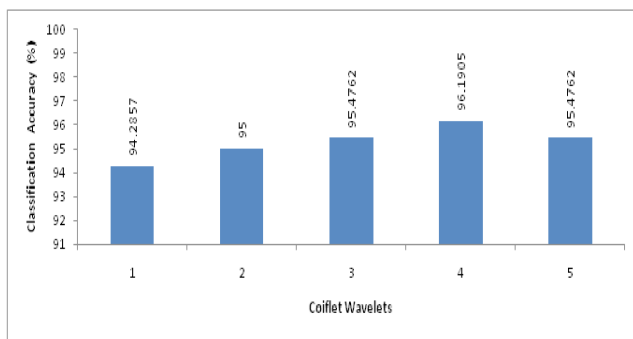


Figure 2. Classification accuracy of coiflet wavelets.

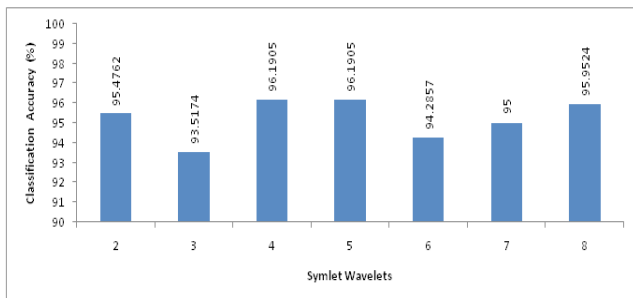


Figure 3. Classification accuracy of symlet wavelets.

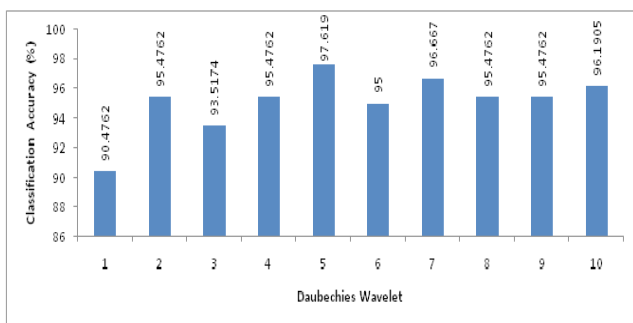


Figure 4. Classification accuracy of daubechies wavelets.

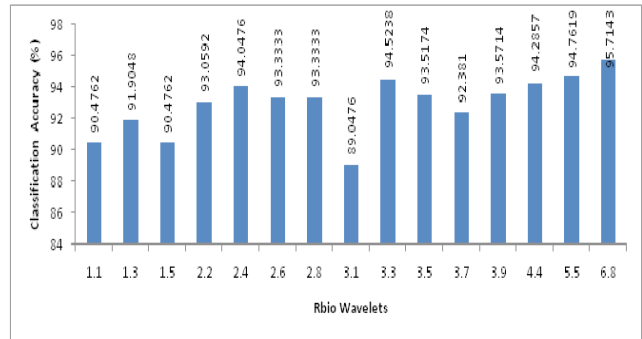


Figure 5. Classification accuracy of rbio wavelets.

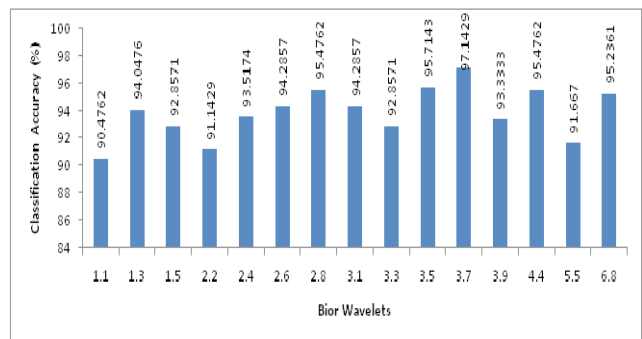


Figure 6. Classification accuracy of bior wavelets.

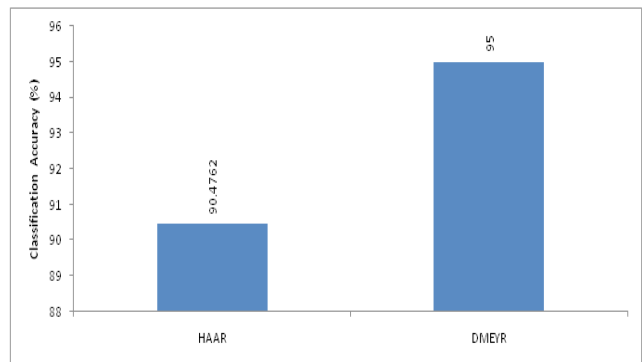


Figure 7. Classification accuracy of HAAR and DMEYR wavelets.

5. Feature Selection using J48 Tree

J48 utilizes a prescient machine-learning model based procedure for information delineation that comprises of branches, leaves, hubs and root to determine principles of order. It executes C4.5 calculation for delivering a pruned or un-pruned C4.5 choice tree. The choice tree considers

many characteristic estimations of the given information and decides the objective of the new specimen. The inner hubs of a choice tree characterize the diverse qualities while the branches between the hubs let us know the conceivable qualities that these traits can have in the watched tests. Accordingly, the last estimation of ward variable is given by the terminal hubs. The choice trees developed by utilizing J48 are connected for further grouping. The building stage and the pruning stage portray the two periods of the J48 choice tree calculation. Amid the building stage, J48 develops choice tree by bridling ideas from data entropy, that is, the tree contains a unitary root hub for the entire preparing set. After each segment, new hubs are added to the choice tree.

For a given set of samples in S , a test attribute called X is chosen for partitioning the set into S_1, S_2, \dots, S_L . Furthermore, new nodes for S are built which are named children and added to the decision tree. The construction of decision tree depends on the test attribute X .

The J48 tree uses each credit of information to manufacture a choice by isolating the information into littler subsets. J48 reviews the standardized data pick up that outcomes from selecting a characteristic for separating the information. The choice is made by utilizing characteristic with the most elevated standardized data pick up. Along these lines, the calculation repeats on the littler subsets and the division technique closes if all cases in a subset have a place with a similar class.

After constructing a Decision tree, the order of attribute selection is followed which is acquired from the tree by examining all the individual attributes and their corresponding values with those observed in the decision tree model. This leads us to a favorable situation where we can ascribe or predict the target values of this new instance. The Information gain (S,A) of a feature A relative to a collection of examples S , is defined as

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{\text{values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v) \quad (1)$$

where $S_v = (\{s \in S \mid A(s) = m\})$.

Entropy is a measure of homogeneity of the set of data and is defined by

$$\text{Entropy}(S) = - \sum_{i=1}^c P_i \log_2 P_i \quad (2)$$

where ' P_i ' is the fraction of ' S ' associated to the class ' i ' and ' c ' is the number of classes.

The other term in the equation above is called expected entropy after S is split using feature A .

While building a tree, J48 sidesteps all the missing qualities, that is, the property estimations of different records can be utilized to foresee the esteem for that thing. Subsequently, the fundamental thought is to part the components into range as indicated by the characteristic qualities for that thing which are available in preparing test¹⁹. At the point when the information turns out to be substantial, the choice tree turns out to be extensive prompting to more incorrectness because of under-fitting or overtraining. Along these lines for more prominent arrangement precision, the trees must be pruned to expel less solid branches.

6. Feature Classification Using Random Forest

Random Forest (RF) is a component grouping technique. It works by building plenty of choice trees amid the preparation time and delivering the class or mean forecast of the comparing trees. Arbitrary choice backwoods correct the choice trees' disservice of over-fitting to the preparation set.

RF can further be distinguished as an accumulation of tree indicators with the end goal that each individual tree depends on the estimations of an arbitrary vector that is examined independently and has indistinguishable circulations for all trees in forest²⁰.

The Random woodland was initially proposed by Leo Breiman in 2000. In Breiman's approach, every tree is framed by selecting indiscriminately, at every hub, a little gathering of info directions to separate the set and moreover by measuring the best division in light of these components among the preparation set. The tree is developed utilizing CART procedure to most extreme size without pruning¹⁶. In random forests, the training algorithm uses a technique called bootstrap aggregating to tree learners. Given a training set $X = x_1, \dots, x_n$ with reactions $Y = y_1, \dots, y_n$, trees are inserted to these samples by bagging repeatedly (B times) by choosing random sample with replacement of training set :

For $b = 1, \dots, B$:

Sample, with replacement, n training examples from X, Y ; named X_b, Y_b .

Train a decision or regression tree f_b on X_b, Y_b .

Following training, predictions for unseen samples x' are possible by averaging the predictions from

all corresponding regression trees on x' :

$$\hat{f} = \frac{1}{B} \sum_{b=1}^B \hat{f}_b(x')$$

Another way of predicting values for unseen samples is by taking the majority vote in the case of decision trees.

Random forests are utilized to grade the precedence of variables in the cases of regression or classification problem in a natural way. Firstly, random forest is fit to the given data in order to measure the variable significance in a data set $(D_n = \{(x_i, y_i)\}_{i=1}^n)$. The out-of-bag error for individual data points is measured during the fitting process and averaged over the forest.

The values of the n -th feature are incorporated among the training data using permutation and the out-of-bag error is again computed on this data set in order to calculate the significance of the n -th feature after training. The found the middle value of contrast in out-of-pack blunder previously, then after the fact the change over all trees is utilized for assessing the significance esteem for the n -th highlight. In the long run, the score is standardized by the standard deviation of these distinctions.

Highlights that create more values for this score are stamped higher than components which deliver lesser qualities. RF is a standout amongst the most dependable calculation and delivers exceedingly exact outcomes. Additionally, it simple to utilize and produces an inward fair gauge of speculation mistake as timberland building advances.

The main reasons for selecting Random forest for the study are listed a follows¹⁹:

1. It's accurate in comparison with other classifiers like PSVM.
2. It's relatively robust to outliers and noise.
3. It's faster than bagging or boosting.
3. It's simple and easily parallelized

7. Results and Discussion

The acoustic signals from the helical gearbox under good condition and different fault conditions (20%, 40%, 60%, 80%, 100% and 150%) were taken. Subsequently feature extraction selection and classification were carried out using DWT, J48 decision tree and Random Forest classifier. Figure 8 gives the maximum classification accuracy of the wavelet families.

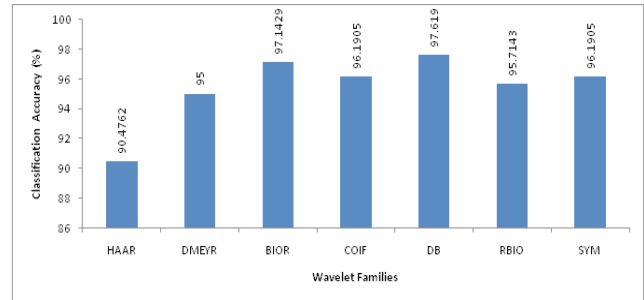


Figure 8. Classification accuracy v/s wavelet families.

1. The Decision tree can be defined as a systematic representation of data that also shows their relative significance.
2. The level of contribution of all wavelet features wasn't discerned to be equal.
3. Out of the 11 wavelet features v1, v2, v3, v4, v5 and v6 were found to play major role in classification of features (Figure 9).

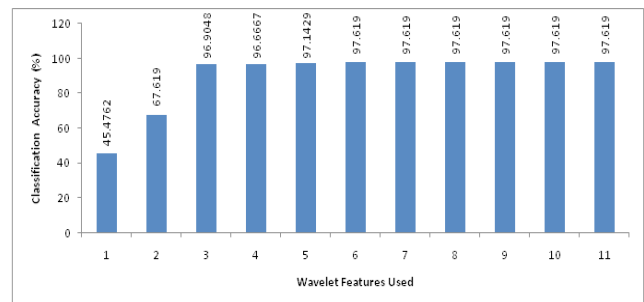


Figure 9. Classification accuracy v/s wavelet features used.

4. These features were selected using the order of precedence in Decision Tree as shown in (Figure 10).
5. The top most wavelet feature v2 has the most dominant effect on the classification and the contribution decreases down the tree.

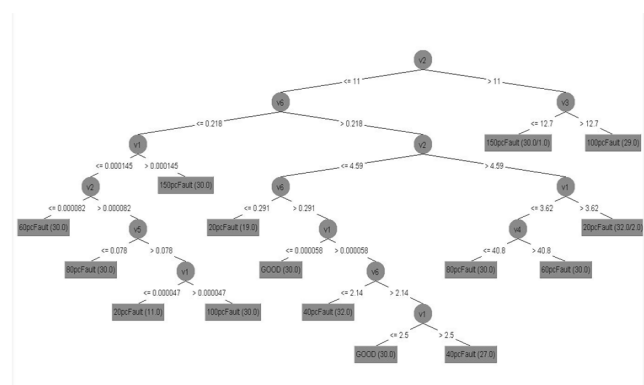


Figure 10. Decision tree.

6. Random Forest Tree was selected since it has been found to be giving the most efficient classification of wavelet features.

7. The classification resulted in number of features (K) be chosen as 2 since it gave the highest accuracy keeping other factors at default values and constant (Figure 11).

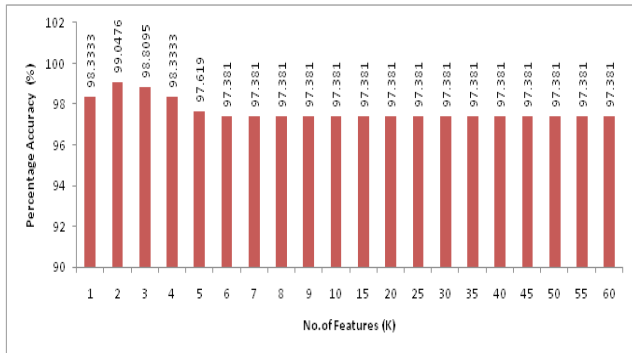


Figure 11. Classification accuracy v/s no. of features (K).

8. Consequently, number of trees (I) having the values 22,24,25,28,29,30,35,40,55 and 60, yielded the same maximum percentage accuracy, keeping the number of feature (K) equal to 2 and other parameters constant. Therefore, the lowest of them all 22 was selected for further operation (Figure 12).

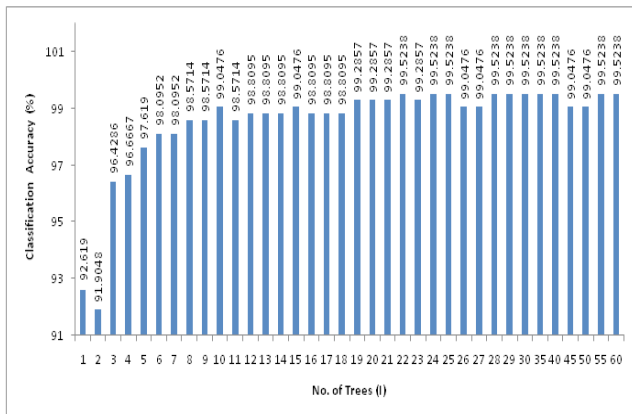


Figure 12. Classification accuracy v/s no. of trees (I).

9. Similarly, varying depth at 0,8,9,10,15,20,25,30,35,40,45,50,55 and 60 (keeping No. of features (K) equal to 2 and No. of trees (I) as 22) the same maximum percentage accuracy resulted, of which the least value 0 is selected (Figure 13).

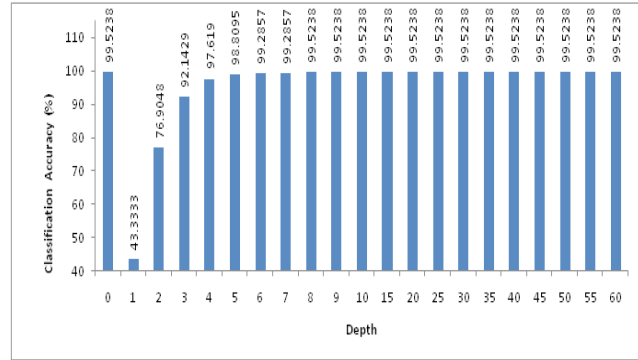


Figure 13. Classification accuracy v/s depth.

10. Finally, number of seeds were varied, keeping No. of features (K) equal to 2, No. of trees (I) equal to 22 and Depth as 0 and maximum percentage accuracy were given by 10,12 and 14, of which the number of seeds were chosen as 10 (Figure 14).

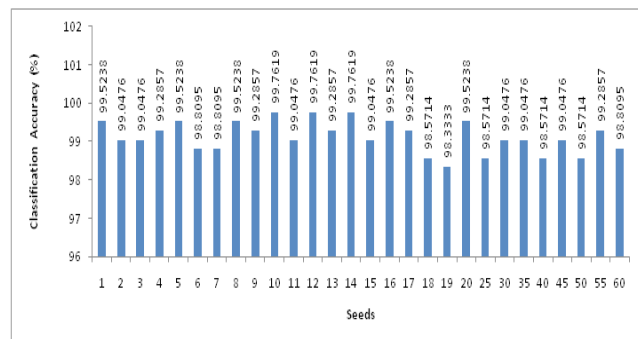


Figure 14. Classification accuracy v/s no. of seed.

Table 1. Confusion Matrix.

Feature	Good	Fault					
		20%	40%	60%	80%	100%	150%
Good	60	0	0	0	0	0	0
20%Fault	0	60	0	0	0	0	0
40%Fault	0	0	60	0	0	0	0
60%Fault	0	0	0	60	0	0	0
80%Fault	0	0	0	0	60	0	0
100%Fault	0	0	0	0	1	59	0
150%Fault	0	0	0	0	0	0	60

11. Table 1 displays the Confusion Matrix which gives an actual insight to the classification process. For features like Good, 20% fault, 40% fault etc. a cent percent classification accuracy has been achieved.

12. Only one instance in the class of 100% fault has been misclassified as a 80% fault class.

13. Correctly Classified Instances 419 99.7619%
 Incorrectly Classified Instances 1 0.2381% Kappa statistic 0.9972

True Positive (TP) rate signifies the amount of items correctly labeled as associated to the positive class that are interpreted as true instances in the same class and ideally should be 1. In the case of class ‘GOOD’, all of the instances are classified correctly and therefore the TP rate is 1. The 20%, 40%, 60%, 80% and 150% are also classified correctly and show TP rate 1 while 100% has a TP rate of 0.983 which means that in most of the validations, the instances were correctly classified. False Positive (FP) means that a given condition has been satisfied, while in reality it has not been satisfied. Pattern recognition and information retrieval consist of precision which is the proportion of recovered instances that are significant and recall that is the proportion of pertinent instances that are recovered. Therefore, both are based on an understanding and estimation of relevance. The measure that integrates precision and recall is harmonic mean of precision and recall is called as the traditional F-measure.

$$\text{Precision} = \frac{\{\{RelevantDocuments\} \cap \{RetrievedDocuments\}\}}{\{\{RetrievedDocuments\}\}}$$

$$\text{Recall} = \frac{\{\{RelevantDocuments\} \cap \{RetrievedDocuments\}\}}{\{\{TotalRelevantDocuments\}\}}$$

$$F = 2 \cdot \frac{\{\{Precision \cdot Recall\}\}}{\{\{Precision + Recall\}\}}$$

The Table 2 gives the detailed accuracy by class.

Table 2. Detailed accuracy by class.

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
Good	1	0	1	1	1	1
20% Fault	1	0	1	1	1	1
40% Fault	1	0	1	1	1	1
60% Fault	1	0	1	1	1	1
80% Fault	1	0.003	0.984	1	0.992	1
100% Fault	0.983	0	1	0.983	0.992	1
150% Fault	1	0	1	1	1	1
Weighted Average	0.0983	0	0.998	0.998	0.998	1

14. The iteration or feature classification in the end yielded a final percentage accuracy of 99.76%.

8. Conclusion

In this study, J48 tree was used for feature selection from which six features out of eleven were selected by taking results from the Decision Tree and Random forest was used for feature classification of acoustic signals from gearbox using wavelet features. Several iterations were conducted on the wavelet features by varying the no. of features, no. of trees, depth and seeds and maximum percentage accuracy was found to be 99.76%. The instances of misclassification of features were minimal in Random Forest as seen on the Confusion Matrix and it proved to be an efficient and precise classifier. Hence, the study concludes that fault diagnostics on the gear box was conducted successfully and various kinds of wavelet features were classified efficiently using random forest which proved to be an easy to use, fast and accurate classifier. Thus, the methodology can be used to provide accurate real time results about the condition of gear teeth.

9. References

- Amarnath M, Krishna IRP. Local Fault Detection in Helical Gears via Vibration and Acoustic Signals using EMD based Statistical Parameter Analysis, Measurement. 2014; 58:154–64
- Sakthivel NR, Indira V, Nair BB, Sugumaran V. Use of Histogram Features for Decision Tree Based Fault Diagnosis of Mono Block Centrifugal Pump, International Journal of Granular Computing, Rough Sets and Intelligent Systems (IJGCRSIS). 2011; 2(1):23–36.
- Wu JD, Liu CH, An Expert System for Fault Diagnosis in Internal Combustion Engines Using Wavelet Packet Transform and Neural Network, Expert Systems with Applications. 2009; 36(3-1):4278–86.
- Yesilurt I, Fault Detection and Location in Gears by the Smoothed Instantaneous Power Spectrum Distribution, NDT&E International. 2003; 36(7):535–42.
- Aharamuthu K, Ayyasamy EP. Application of Discrete Wavelet Transform and Zhao-Atlas-Marks Transforms in Non-Stationary Gear Fault Diagnosis, Journal of Mechanical Science and Technology. 2013; 27(3):641–47.
- Vrhel MJ, Lee C, Unser M, Rapid Computation of the Continuous Wavelet Transform by Oblique Projections, IEEE Transactions on Signal Processing. 1997; 45(4):891–900.
- Szabo L, Dobai JB, Biro KA, Discrete Wavelet Transform based Rotor Faults Detection Method for Induction Machines, Department of Electrical Machines, Marketing and Management, Technical University of Cluj, 2008, p 1–12.
- Quinlin JR. Induction of Decision Trees, Machine Learn-

- ing, Kluwer Academic Publishers, Boston - Manufactured in The Netherlands. 1986; 97(1-2):321–27.
9. Garg R, Mittal S. Optimization by Genetic Algorithm, *International Journal of Advanced Research in Computer Science and Software Engineering*. 2014;4(4):111–23.
 10. Banerjee A. Impact of Principal Component Analysis in the Application of Image Processing, *International Journal of Advanced Research in Computer Science and Software Engineering*. 2012; 61(3):611–22.
 11. Saravanan N, Siddabattuni VNSK, Ramachandran KI. A Comparative Study on Classification of Features by SVM and PSVM Extracted using Morlet Wavelet for Fault Diagnosis of Spur Bevel Gear Box, *Expert Systems with Applications*. 2008; 35(1):1351–66.
 12. Durgesh K, Srivastava S, Bhambhu L. Data Classification Using Support Vector Machine, *Journal of Theoretical and Applied Information Technology*. 2009; 12(1):1–3.
 13. Sakthivel NR, Sugumaran V, Nair BB. Automatic Rule Learning using Roughset for Fuzzy Classifier in Fault Categorization of Centrifugal Pump, *International Journal of Applied Soft Computing*. 2012; 12(1):196–203.
 14. Pyke SW, Sheridan PM, Logistic Regression Analysis of Graduate Student Retention, *The Canadian Journal of Higher Education*. 1993; 23(2):21–76.
 15. Chharia A, Gupta RK. Enhancing Naïve Bayes Performance with Modified Absolute Discount Smoothing Method in Spam Classification, *International Journal of Advanced Research in Computer Science and Software Engineering*. 2013; 3(3):1–6.
 16. Biau G. Analysis of a Random Forests Model, *Journal of Machine Learning Research*. 2012; 13(1):1063–95.
 17. Saravanan N, Ramachandran KI. Incipient Gear Box Fault Diagnosis using Discrete Wavelet Transform (DWT) for Feature Extraction and Classification using Artificial Neural Network (ANN), *Expert Systems with Applications*. 2010 Jun; 37(6):4168–81.
 18. Ghate VN, Dudul SV. SVM based Fault Classification of Three Phase Induction Motor, *Indian Journal of Science and Technology*. 2009 May; 2(4):32–35.
 19. Patil TR, Sherekar SS. Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification, *International Journal of Computer Science and Applications*. 2013 Apr; 6(2):256–61.
 20. Breiman L, Random Forests. Statistics Department, University of California, 2001.
 21. Jain R, Abraham A. A Comparative Study of Fuzzy Classification Methods on Breast Cancer Data, *Australasian Physics and Engineering Sciences in Medicine*. 2004 Dec; 27(4):213–18.