ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, Vol 9(33), DOI: 10.17485/ijst/2016/v9i33/101325, September 2016 ISSN (Online) : 0974-5645

Fault Diagnosis of Bearings using Vibration Signals and Wavelets

V. Dhanush Abhijit[1*], V. Sugumaran[1] and K. I. Ramachandran[2]

[1]School of Mechanical and Building Science, SMBS, VIT University, Chennai - 600237, Tamil Nadu, India; venugopaldhanush@gmail.com, v_sugu@yahoo.com

[2]CEN, Amrita School of Engineering, Ettimadai, Coimbatore – 600127, Tamil Nadu, India; ki_ram@cb.amrita.edu

Abstract

Objectives: Being widely used in most of the industrial machineries, bearings are subjected to wear and tear. Failure of bearings can incur heavy losses in the industries. In order to prevent such mishaps during operation, it is necessary to subject the bearings to a suitable fault diagnosis technique. **Methods/Statistical Analysis**: Vibration analysis is performed to detect the fault in bearings. For the fault analysis, vibration signals were taken for good, inner race defect, outer race defect and combination of these defects. Since vibration signals are complex and the defect related signature is buried deep within the noise and high frequency resonance, simple signal processing cannot be used for effectively detecting bearing fault. In this paper, discrete wavelets transform were used to detect bearing faults. For wavelet and feature selection, J48 decision tree algorithm was used. For feature classification, Best First Tree (BFT) algorithm was used. **Findings:** The experimental results indicate biorthogonal wavelets show maximum successful bearing fault detection rate. The classification accuracy was calculated and found to be 96.25%. This result is further refined to get better classification accuracy and the final result was found to be 98%. **Application/Improvements:** This can be considered to be a part of a preventive maintenance method in order to avoid mishaps in industries. The classification accuracy can be further improved using different algorithms.

Keywords: Biorthogonal, Decision Tree, Fault Diagnosis, Feature Selection, Vibration Signals, Wavelet Selection, Wavelet Transforms

1. Introduction

Bearings are machine elements that provide free rotation movement around a fixed axis and constrain the motion to only desired direction. The bearings are the most important components in rotary machines. The life of a rolling element bearing is determined by exposing temperature, carrying loads, maintenance frequency, proper lubrication, handling, installation etc. The overall performance is affected by its carrying capacity and reliability.

The rolling element wears out easily due to metal to metal contact. These wears create faults in the outer and inner races. It is also the vulnerable part of a machine because it works under heavy load and high rotational speed. The breakdown of bearings can cause breakdown of machines which may even result in heavy financial losses or human casualties. Due to this very reason, monitoring the condition of bearings is essential for early warning.

Vibration analysis is a suitable method for fault diagnosis. Using an accelerometer, vibration signals can be measured directly. Readings of good and fault conditions are compared to analyze the condition of bearings. Defects in bearings cause variation in the frequency at which it operates and this signal is modulated by natural frequency of the bearings. Since, the signature of defective bearing is spread across a wide frequency band, it is difficult to detect it.

As per statistics, about 90% of all types of faults in rolling bearings are either an Inner Race Fault or Outer Race Fault[1]. In this paper, the Inner Race Fault, Outer Race Fault and combination of faults are considered for fault diagnosis.

*Author for correspondence

In order to detect and recognize the fault signal, signal analysis methods were used. In earlier studies, Fourier transforms[2] were the dominating methods in signal analysis. The signal has to be periodic and stationary for using Fourier transforms. Since, the vibration signals may not always be stationary, Fourier method is not reliable in all situations. Hence, methods were introduced to simultaneously generate both time and frequency information. Among the time-frequency analysis methods, wavelets are the most widespread tools in signal analysis.

After the signal analysis, three steps are involved which are: Feature extraction, feature selection and feature classification. In this paper, wavelet transforms[3] was considered for feature extraction. For feature selection, decision tree[4] technique was used as they were easy to understand and easy to use. For feature classification, various classifiers are used by researchers. They are Logistic regression[5], Artificial Neural Network (ANN)[6], fuzzy logic[7], decision tree, Support Vector Machine (SVM) and Proximal Support Vector Machine (PSVM)[8–10].

In[11] investigated faults in ball bearing using wavelet transforms and Adaptive Neural-Fuzzy Interface System (ANFIS) as classifier[11]. In[12] did fault diagnosis of bearing using Wavelet Analysis (WA) and Fast Fourier Transform (FFT) and found that wavelet analysis provide better resolution compared to FFT in low frequency range[12]. In used wavelet transforms to extract features in order to compare the artificial intelligence techniques (ANN, SVM and logistic regression) in rolling element bearings. In[13] presented a method for designing a new wavelet using Continuous Wavelet Transform however the classification accuracy was only 83.25%[13]. In[14] used Discrete Wavelet Transform (DWT) for fault diagnosis of centrifugal pump using j48 algorithm[14]. In[15] compared the use of DWT over FFT for detection of ball bearing race faults for single and multiple point defects in inner race, outer race and combination of these faults[15]. In[16] conducted a comparative study of decision tree classifier and Best First Tree classifier for fault diagnosis of hydraulic brake system[16].

In this paper, Discrete Wavelet Transform was used for feature extraction. J48 decision tree algorithm was used for wavelet selection and feature selection. Best First Tree classifier for feature classification.

2. Experimental Studies

When bearing is in faulty condition, the main objective is to segregate the faults into Outer Race Fault, Inner Race Fault or combination of faults. This paper focuses the use of wavelets and decision tree for fault diagnosis of bearings. The experimental setup and procedure are as follows:

2.1 Experimental Setup

Two roller bearings are attached to a short shaft of diameter 30 mm. The short shaft is connected to the variable speed DC motor of 0.5 hp with a rated rpm of 3000 rpm through a flexible coupling. The flexible coupling is used to reduce the transmission of vibration and effects of misalignment. The bearing close to the motor is a brand new bearing so that it can be assumed to be free of defect. The second bearing is the bearing under test. The piezoelectric accelerometer is mounted on top of the bearing housing using direct adhesive techniques.

The accelerometer is connected to a signal conditioning unit (DACTRAN FFT analyzer). It consists of a charge amplifier and Analog-Digital Converter (ADC). The signal in digital form is obtained by computer using the software RT Pro-series. The signal is stored in the memory and can be processed to extract the features. The experimental setup is shown in Figure 1.

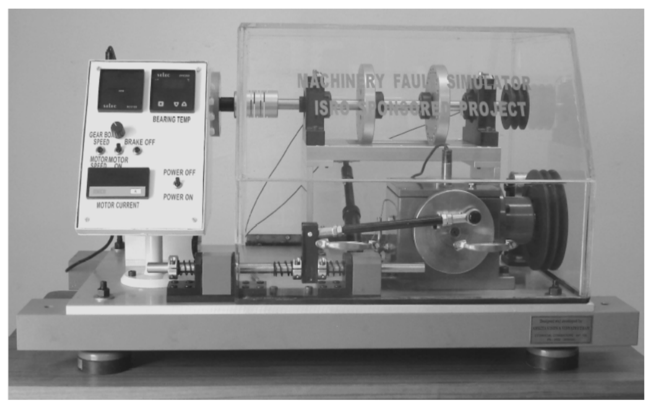

Figure 1. Bearing test setup.

2.2 Experiment Procedure

Four bearings were used in the experiment. One was a new bearing assumed to have no defects and in the other three bearings, defects were created. EDM method was used to create the defects in order to keep the defect size under control. A cut of 0.525 mm wide and 0.827 mm deep cut was done on the inner race of one bearing and a cut of 0.652 mm wide and 0.981 mm deep cut was done on the outer race of another bearing. These two defects were combined in the third bearing.

After running the bearing for some time, vibration signals were taken using the mounted piezoelectric accelerometer. For all speeds and conditions, the sampling frequency was 12000 Hz and length of sample was 8192. Sample length of 10000 was chosen however for wavelet feature extraction, the number of samples should be 2^n. The nearest 2^n is 8192. Hence, 8192 was chosen as the sample length. This test was repeated by varying the motor speed for 700 rpm, 800 rpm and 900 rpm.

3. Feature Extraction

The vibration signals were analyzed to perform fault diagnosis. The obtained signal was time-domain signal. This time-domain signal was converted into time-frequency-domain data by using Discrete Wavelet Transform (DWT). Wavelet decomposition was performed using DWT on vibration signals. The decomposition gives trend and details. The trends was again decomposed into next level trend and details. The trends of previous levels were subsequently decomposed and many levels of details were obtained. The length of the signal was 8192 (2^{13}) and thus, 13 levels of decomposition were possible. At each level, the detail co-efficient were used to compute energy content using the following formula.

$$V_i = \sum_{i=1}^{n} Xi2$$

Where $x_{i\,=}$ details coefficients.

N = number of details coefficients.

Then the features were defined as the energy content at each level. The feature vector is defined as:

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

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

$v_1, v_2, v_3 \ldots$ are energy content at given level

Various families of wavelets were considered 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

Using Discrete Wavelet Transform (DWT), the features from time-domain signal were extracted. 13 vectors were formed which are v_1, $v_{2,}$ $v_{3,}$ $v_{4\ldots\ldots}$ v_{13}. Before selecting the features that contribute to the maximum classification accuracy, the wavelet has to be selected. The classification accuracy of 7 families of wavelets and their child wavelets were found using J48 decision tree. The classification accuracy were calculated and compared.

As shown in Figure 2 – Figure 7. Rbio 3.1 (95.75%), db 10 (94.25%), coil 5 (93.75%), sym 8 (91.25%), Bior 3.1 (96.25%), dmey (96.25%) and haar (92.25%) were maximum classification accuracy obtained for each wavelet family. The wavelets with highest classification accuracy from each family were taken and compared. From Figure 8, it is clear that the bior 3.1 and dmey has the maximum classification accuracy. Among them, bior 3.1 alone was chosen for further calculations.

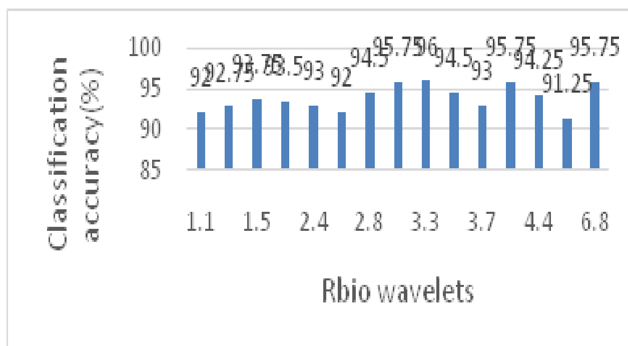

Figure 2. Classification accuracy of rbio wavelets.

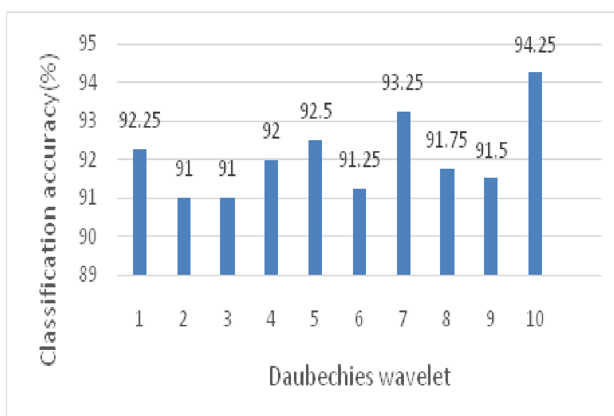

Figure 3. Classification accuracy of Daubechies wavelet.

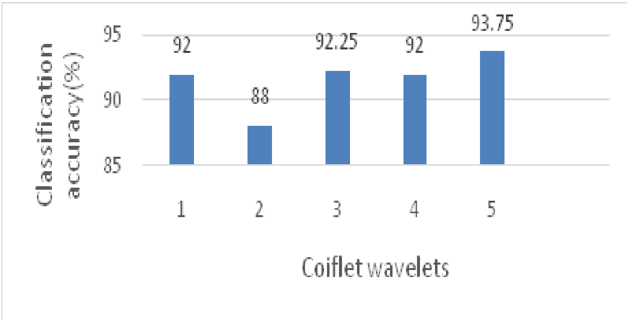

Figure 4. Classification accuracy of Coif lets.

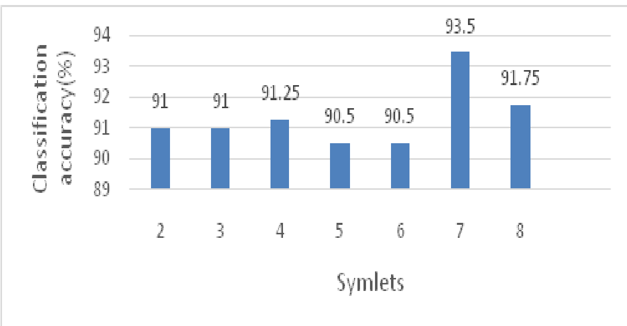

Figure 5. Classification accuracy of Symlets.

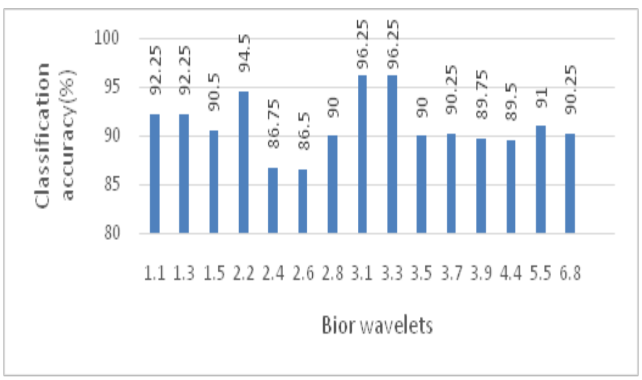

Figure 6. Classification accuracy of Bior wavelets.

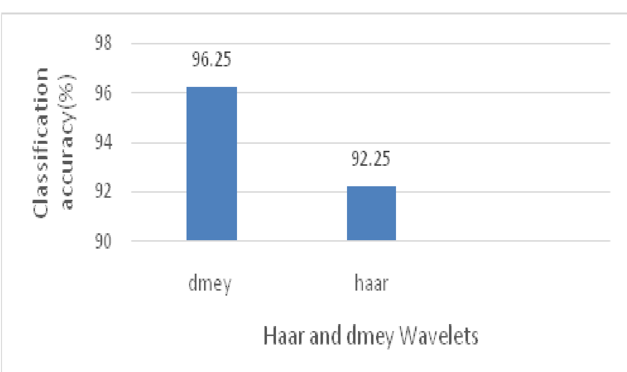

Figure 7. Classification accuracy of dmey and haar wavelets.

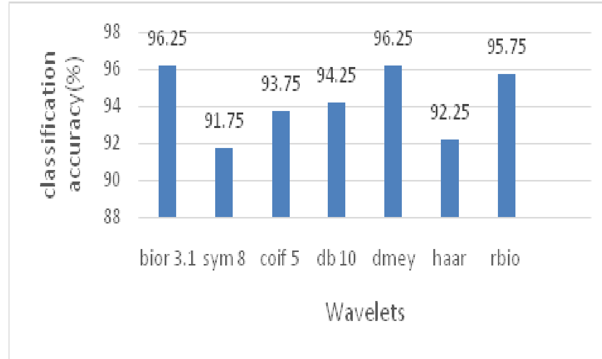

Figure 8. Classification accuracy of wavelets.

5. Feature Selection

Among the 13 features that were extracted, all of them may not contribute to the classification accuracy. Some features may have no effect or reduce the classification accuracy. These features are called irrelevant features. Feature selection was carried out to take relevant features from the 13 features. This helps in ignoring the irrelevant features. The feature selection was carried out using decision tree algorithm.

Decision tree is a knowledge representation method used to represent classification rules.

J48 decision tree is used for feature selection. In a J48 decision tree, there is a root and a number of branches, nodes and leaves. The tree starts from a single node which is called the root. Root is the most significant feature. Branch is the chain that connects nodes from root to leaves. The bottom end of every branch is a leaf node which represent class label. Each node in the tree represents a feature and the occurrence of the feature in tree represent the importance of the feature in this fault diagnosis process.

Following the footsteps[8] of the feature selection was carried out. The features that contribute for the feature classification were selected from the tree Figure 9.

From the decision tree Figure 9, it is clear that v_3, v_4, v_1 and v_{13} are the features that contribute to the feature classification. The classification accuracy according to the number of features are calculated. At first, v_3 alone was used to find classification accuracy and it was found as 94.25%. After that, the next feature from the decision tree is taken *i.e,* v_4. The classification accuracy of both were calculated and found out to be 96.5%. This process was carried out increasing the number of features. When all the features *i.e*. v_3, v_4, v_1, v_{13} from the decision tree is used, the classification accuracy is 97.25%. On increasing the number of features the accuracy was found to be reducing Figure 10.

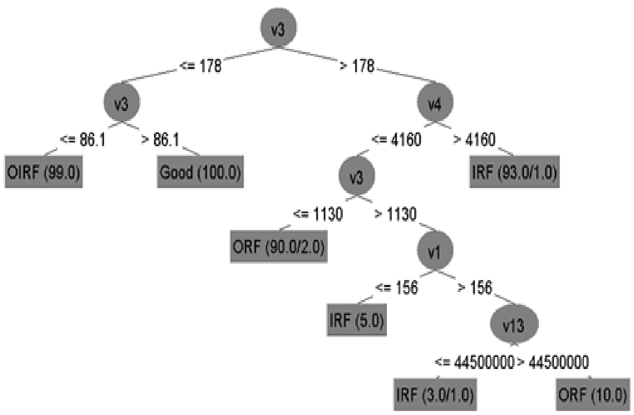

Figure 9. Decision tree.

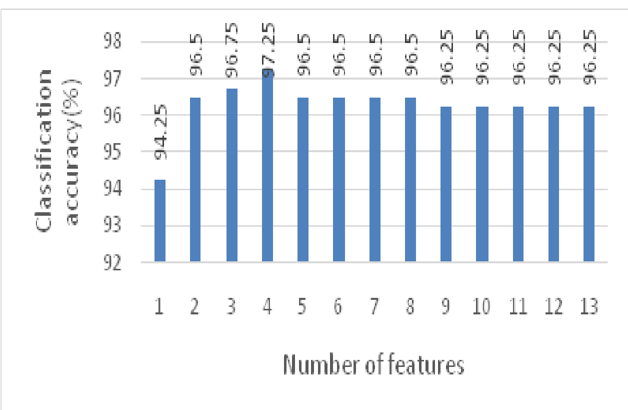

Figure 10. Effect on classification accuracy with number of features.

6. Feature Classification

Feature classification is the process of pruning the decision tree by removing the features that has less discriminating ability to increase the accuracy and robustness. Pruning was carried out to avoid over fitting. Overfitting occurs in case of a complex model or when the model triesw to memorize training data rather than learning to generalize from trend. An algorithm is said to be overfit when it is more accurate in fitting known data and less accurate to predict new data. A decision tree is pruned to get a tree that generalizes better to independent test data.

In this paper, the Best First Tree (BFT) algorithm was used for feature classification. In BFT, classification is carried out starting from the root to the terminal node, by testing the features and moving down the branches according to the value of features. The first node to be split in this tree is the best node which is actually responsible for maximum reduction in irrelevant attributes in the decision tree. The node impurity is measured to find the best node using the splitting criteria. After finding the best attribute, the next best node to be expanded is found out. The tree is stopped when fixed number is specified, using the stopping criteria. This enables pruning methods by choosing the fixed number of expansions.

Pruning is done by varying the values of certain parameters of the decision tree. In the Best First Tree (BFT) algorithm, the parameters that control the pruning of the tree are: Minimum number of instances (M), Number of folds in internal cross-validation (N), seed (S) and percentage of Training set size (C). Among these parameters, the minimum number of instances is more important compared to the rest as it contributes more in increasing the classification accuracy. Minimum number of instances depend upon the data set. In this paper, it ranges from 1 to 100. The minimum number of instances (m) value is changed from 5 to 100 with the increase of 5 in each step and it was found that classification accuracy is maximum at m = 5. On further calculating the classification accuracy when m ranges from 1 to 5, it was noticed that the maximum classification accuracy is obtained at m = 2 Figure 11.

The minimum number of instances was fixed and number of folds in internal cross- validation (N) is varied stating from 2. The number of times the cross-checking with different values are performed is the number of folds. The change in classification accuracy with respect to the number of folds is calculated and plotted Figure 12. At N = 3, the classification accuracy was found to be maximum compared to the rest of the N values. The calculation was stopped at N = 20 because the classification accuracy shows no variation after N = 5.

Values of M and N were fixed and seed value was varied. Seeds are randomly selected instances from the dataset. The seed value was increased by steps of 1 from 1 and corresponding classification accuracy was calculated and plotted Figure 13. It was found that when seed is 7 classification accuracy is maximum.

The last parameter to be varied was the percentage of training set of data (C) which varies from 0 to 1. The training set of data was varied from 0.1 to 1 at steps of 0.1 Figure 14. At C = 1, the maximum classification accuracy of 98% was calculated.

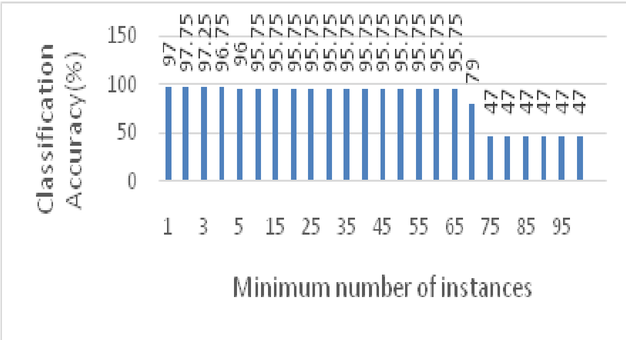

Figure 11. Effect of number of instances (m) on classification accuracy.

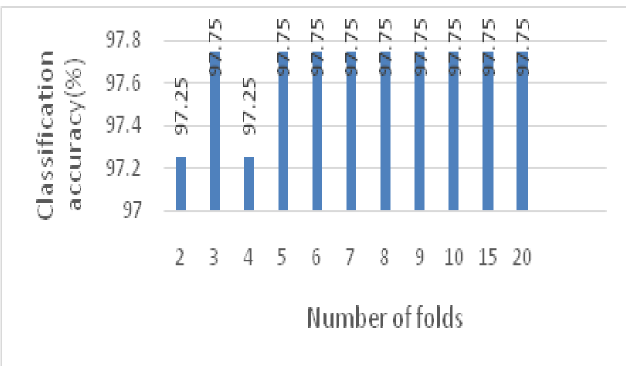

Figure 12. Effect of number of folds on classification accuracy.

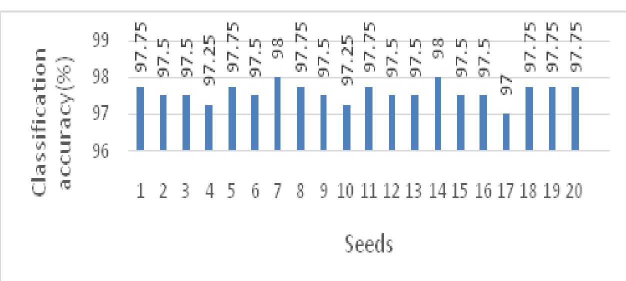

Figure 13. Effect of seeds on classification accuracy.

7. Results and Discussions

Discrete Wavelet Transforms (DWT) was used to complete feature extraction. Various families of wavelets are used for feature extraction. Biro 3.1 was selected among the families of wavelet and the rest of the wavelets are ignored for further process.

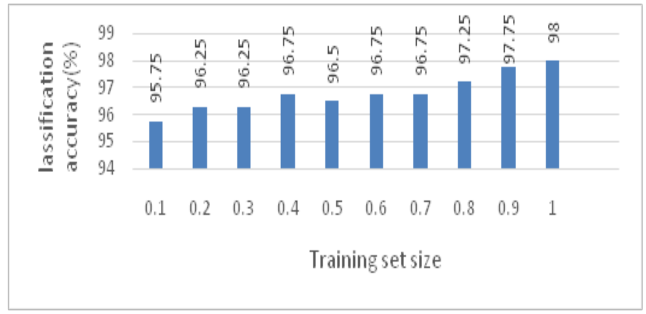

Figure 14. Effect of training set size on classification accuracy.

The decision tree for biro 3.1 is generated and the features contributing to the maximum classification accuracy are v_3, v_4, v_1 and v_{13}. It is confirmed that these 4 features are the relevant features by comparing the classification accuracy of number of features. Once the features are selected using J48 algorithm, feature classification was carried out. The features were trained and tested using 10-fold cross validation.

The stratified cross-validation summary is as follows:

Correctly Classified Instances	392	98%
Incorrectly Classified Instances	8	2%
Kappa statistic	0.9733	
Mean absolute error	0.0142	
Root mean squared error	0.0965	
Relative absolute error	3.7768%	
Root relative squared error	22.2893%	
Total Number of Instances	400	

Table 1. Confusion matrix

	GOOD	ORF	IRF	OIRF
GOOD	99	1	0	0
ORF	1	97	2	0
IRF	0	3	97	0
OIRF	0	1	0	99

The classification accuracy of the BFT algorithm is in the form of confusion matrix Table 1. The correctly classified instances are shown as the diagonal elements of the configuration matrix. First row represents good readings *i.e.*, readings taken from a bearing with no fault. Among these readings, one reading is wrongly classified under Outer Race Fault (ORF). Second row represents ORF. Among these 100 data points, 97 are classified under ORF and 3 are misclassified. Among the misclassified data

Table 2. Detailed accuracy

Class	TP rate	FP rate	Precision	Recall	F-Measure	ROC area
Good	0.99	0.003	0.99	0.99	0.99	0.995
ORF	0.97	0.017	0.951	0.97	0.96	0.977
IRF	0.97	0.07	0.98	0.97	0.975	0.989
OIRF	0.99	0	1	0.99	0.995	0.994
Weighted average	0.98	0.007	0.98	0.98	0.98	0.989

point, 2 are under Inner Race Fault (IRF) and 1 is under Good. Third row represents IRF. Among them, 97 are correctly classified under IRF, however 3 of the data points are misclassified under ORF. Fourth and last row represents Outer and Inner Race Fault (OIRF). 99 are classified correctly under OIRF and one is misclassified under ORF. The detailed accuracy by class is as follows Table 2.

True Positive (TP) rate should be 1 and False Positive (FP) should be 0 for ideal cases. For Good condition, the TP rate is found to be 0.99 which means out of 100, 99 are correctly labeled to the good class. From the given table, it is noticed that all classes have a TP rate closer to 1. FP rate depicts the amount of the unfulfilled conditions that are indicated as full filled. The precision and recall are a measure of relevance. They should be 1 for ideal cases. F-measure is the harmonic mean of precision and recall. ROC area is the area under the curve plotted with TP rate *vs.* FP rate. The percentage of randomly drawn pairs for which classification is true is called the ROC area. The ROC area should be 1 for ideal cases. From the Table 1, it is clear that the precision, recall and ROC area are approximately equal to 1.

8. Conclusion

The Inner Race Faults, Outer Race Faults and combination of Inner Race and Outer Race Faults were simulated on a rolling bearing. The vibration signals were obtained using accelerometer. The time domain signal was converted into time-frequency domain using wavelet transforms. Among the wavelet families that were used, biro 3.1 provided the maximum classification accuracy for the fault diagnosis. J48 decision tree was used for wavelet selection and feature selection. Best First Tree (BFT) algorithm was used for feature classification. It provides a classification accuracy of 98%. Studies are limited as experiment was carried out in simulated bearing faults.

The main objective of fault diagnosis is to detect the presence of fault in bearing during a fast process. Depending on the requirement and company policies, maintenance can be carried out. Due to high computation speed and easiness, this method is reliable for fault diagnosis of bearings. Further studies are possible by conducting the experiment in real time systems.

9. Reference

1. Bently D. Predictive maintenance through the monitoring and diagnostics of rolling element bearings. Applications Note, ANO44. Bently Nevada Co; 1989. p. 2–8.
2. Peng ZK , Chu FL. Application of wavelet transform in machine condition monitoring and fault diagnostics: A review with bibliography. Mechanical Systems and Signal Processing. 2004 Mar; 18(2):199–21.
3. Nikolaou NG, Antoniadis IA. Rolling element bearing fault diagnosis using wavelet packets. NDT&E International. 2002 Apr; 35(3):197–205.
4. Sakthivel NR, Sugumaran V, Babudevasenapati S. Vibration based fault diagnosis of monoblock centrifugal pump using decision tree. International Journal of Expert Systems with Applications. 2010 Jun; 37(6):4040–9.
5. Pandya DH, Upadhyay SH, SHarsha SP. Fault diagnosis of rolling elements by using multinomial logistic regression and wavelet packet transform. **Methodologies And Application.** 2014 Feb; 18(2):255–66.

6. Rajakarunakaran S, Venkumar P, Devaraj D, Rao KSP. Artificial Neural Network approach for fault detection in rotary system. Applied Soft Computing. 2008 Jan; 8(1):740–8.
7. Sakthivel NR, Sugumaran V, Nair B. Comparison of decision tree-fuzzy and rough set-fuzzy methods for fault categorization of mono-block centrifugal pump. Mechanical Systems and Signal Processing. 2010 Aug; 24(6):1887–906.
8. Sugumaran V, Muralidharan V,. Ramachandran KI. Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing. Mechanical Systems and Signal Processing. 2007 Feb; 21(2):930–42.
9. Muralidharan V, Sugumaran V. A comparative study of Naive Bayes classifier and Bayesnet classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. Journal of Applied Soft Computing. 2012 Aug; 12(8):2023–9.
10. Sugumaran V, Kumar TAR, Amarnath M Hemanta Kumar. Fault diagnosis of bearings through vibration signal using Bayes classifiers. International Journal for Computer Aided Engineering and Technology. 2014 Jan; 6(1):14–28.
11. Lou X, Loparo KA. Bearing fault diagnosis based on wavelet transform and fuzzy inference. Mechanical Systems and Signal Processing. 2004 Sep; 18(7):1077–95.
12. Tse PW, Peng YH, Yam R. Wavelet analysis and envelope detection for rolling element bearing fault diagnosis - Their effectiveness and flexibilities. Journal of Vibration and Acoustic. 2001 Mar; 123(3):303–10.
13. Manju BR, Rajan AR, Sugumaran V. Wavelet design for fault diagnosis of Roller Bearings using Countinous Wavelet Transforms. IJMET. 2010 Jul–Aug; 1(1):38–48.
14. Muralidharan V, Sugumaran V. Selection of Discrete Wavelets for fault diagnosis of monoblock centrifugal pump using the J48 algorithm. Applied Artificial Intelligence. 2013 Jan; 27(1):1–19.
15. Prabhakar S, Mohanty AR, Sekhar AS. Application of Discrete Wavelet Transform for detection of ball bearing race faults. Tribology International . 2002 Dec; 35(12):793–800.
16. Jagadeeshwaran R, Sugumaran V. Comparative study of decision tree classifier and Best First Tree classifier for fault diagnosis of automobile hydraulic brake system using statistical features. Measurement. 2013 Nov; 46(9):3247–60.