

# Artificial Neural Network Based Fault Diagnostics of Rolling Element bearings using Continuous Wavelet Transform

R. Zaeri, B. Attaran, A. Ghanbarzadeh, S. Moradi

**Abstract**—Any industry needs an efficient predictive plan in order to optimize the management of resources and improve the economy of the plant by reducing unnecessary costs and increasing the level of safety. A great percentage of breakdowns in productive processes are caused by bearings. This paper presents a methodology for fault diagnosis of ball bearings based on continuous wavelet transform (CWT) and artificial neural network (ANN). Three wavelet selection criteria Maximum Energy, Minimum Shannon Entropy, and Maximum Energy to Shannon Entropy ratio are used and compared to select an appropriate wavelet to extract statistical features. Total 15 feature set and 87 mother wavelet candidates were studied, and results show that complex morlet 1-1 has a best diagnosis performance based on minimum shannon entropy than the other mother wavelets and criteria. Also results show the potential application of proposed methodology with ANN for the development of on-line fault diagnosis systems for machine condition.

## I. INTRODUCTION

ROLLING element bearings are widely used in various types of machines ranging from simple electric fans to complex manufacturing facilities. Bearing faults, in fact, are a common cause of machinery failures. Therefore, an effective bearing fault diagnostic technique is critically needed for a wide array of industries for early detection of bearing defects so as to prevent machinery performance degradation and malfunction. Several methods have been proposed in the literature for bearing fault detection. To inspect raw vibration signals, a wide variety of techniques have been introduced that may be categorized into two main groups: classic signal processing and intelligent systems. To make mention of a few, FFT, Wigner-Ville distribution, wavelet are classic signal processing methods [1]-[3]. ANN-based and SVM could be classified as intelligent systems [4]-[6]. Currently, industrial applications of intelligent monitoring systems have been increased by the progress of intelligent systems. Rafiee et al. are applied new wavelet selection method and studied 324 mother wavelet and results shows that db44 has most similar shape across both gear and bearing vibration signals but it is not the proper function for

all wavelet-based processing [7]. Kankar et al. have presented two wavelet selection criteria *Maximum Energy to Shannon Entropy ratio* and *Maximum Relative Wavelet Energy* to select an appropriate wavelet for feature extraction. They fed four statistical features as input to SVM, ANN, SOM classifiers and results showed that SVM identified the fault categories of rolling element bearing more accurately for both Meyer wavelet and Complex Morlet wavelet and has a better diagnosis performance as compared to the ANN and SOM [8].

Wavelet transform (WT), capable of processing stationary and non-stationary signals in time and frequency domains simultaneously, was used for feature extraction [9]. Previous research on wavelet transform, focused on application and advantage of WT when applied on fault diagnosis. WT can be mainly divided into discrete and continuous forms. The former is faster with lower CPU time, but continuous forms are more efficient since there is good resolution throughout the signals. In this paper a methodology is proposed for selection of most appropriate wavelet and to determine scale corresponding to characteristic defect frequency based on wavelet selection criterion. These raw signals are divided into  $2^7$  sub-signals, i.e. 128 scales in seventh level of decomposition to convert the complex vibration signals into simplified signals with more resolution. 87 different wavelets are considered with each  $2^7$  sub-signals, i.e. 128 scales. Three wavelet selection criteria *Maximum Wavelet Energy*, *Minimum Shannon Entropy* and *Maximum Energy to Shannon Entropy ratio* are used and compared to select an appropriate wavelet for feature extraction. Statistical features are calculated from continuous wavelet coefficients and are fed as input to artificial neural network (ANN). The results showed that the proposed methodology can extract useful features from the original data. A comparison between the present method and the other methods showed that the presented method is remarkably efficient in on-line fault diagnosis systems for machine condition. Also Flowchart of the presented intelligent fault diagnosis system is shown in fig.1. This is mention that results of this work in comparison with kankar et al. [8] work has been more general, better and also results are obtained for 87 mother wavelet using only back propagation ANN as a classifier but with 99.68% accuracy.

## II. THEORY OF ANN

### A. Review of machine learning techniques

Machine learning is an approach of using examples (data) to synthesize programs. In the particular case when

---

Manuscript received October 3, 2011.

R. Zaeri. Mechanical Engineering Department, University of Shahid Chamran, Ahvaz, Iran (corresponding author to provide phone: +986115524843; e-mail: Reza.Zayeri@ yahoo.com).

B. Attaran. Mechanical Engineering Department, University of Shahid Chamran, Ahvaz, Iran (e-mail: Reza.Zayeri@ yahoo.com).

A. Ghanbarzadeh. Mechanical Engineering Department, University of Shahid Chamran, Ahvaz, Iran (e-mail: Ghanbarzadeh.A@scu.ac.ir).

S. Moradi. Mechanical Engineering Department, University of Shahid Chamran, Ahvaz, Iran (e-mail: moradis@scu.ac.ir).

the examples are input/output pairs, it is called supervised learning. In a case, where there are no output values and the learning task is to gain some understanding of the process that generated the data, this type of learning is said to be unsupervised. In the present study, the back propagation (BP) algorithm is considered. Pattern recognition and classification using machine learning techniques are described here [10].

### B. Supervised Learning

Artificial neural network is an interconnected group of artificial neurons. These neurons use a mathematical or computational model for information processing. ANN is an adaptive system that changes its structure based on information that flows through the network [10]. A single neuron consists of synapses, adder and activation function. Bias is an external parameter of neural network. Model of a neuron can be represented by following mathematical model.

$$y_k = \phi\left(\sum_{i=1}^p w_{ki} x_i + w_{k0}\right) \quad (1)$$

Input vector comprises of 'p' inputs multiplied by their respective synaptic weights, and sum off all weighted inputs. A threshold (bias) is used with constant input. Activation function converts general output into a limited range of output. Intelligence of neural network lies in the weights between neurons. BP algorithm is used as learning algorithm for calculating synaptic weights.

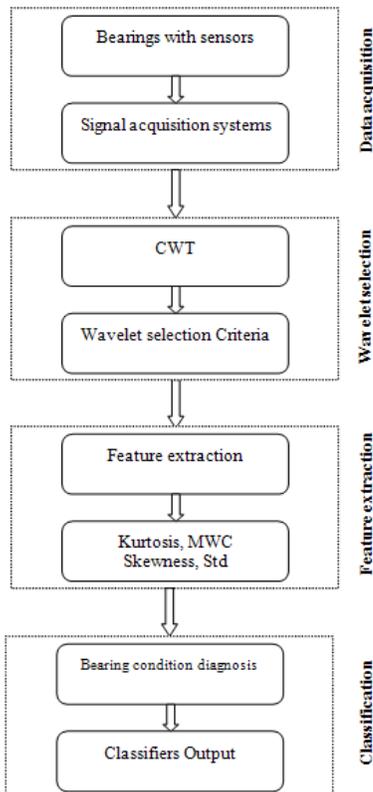


Fig. 1. Flowchart of the presented intelligent fault diagnosis system

### III. EXPERIMENTAL SETUP AND DATA ACQUISITION

The bearing vibration signal of four types of bearing conditions were extracted from Case Western Reserve University bearing test data center [11]. The ball bearings used in the experiment are installed in a motor driven mechanical system as shown in fig. 2.

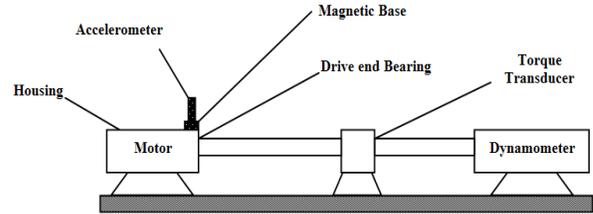


Fig. 2. Schematic of Experiment System

The test stand is made up of a 2 hp Reliance Electric motor, a torque transducer-encoder, a dynamometer, and control electronics. Vibration signals from the bearings were acquired using an accelerometer that was attached to motor housing at the drive end of the motor with a magnetic base.

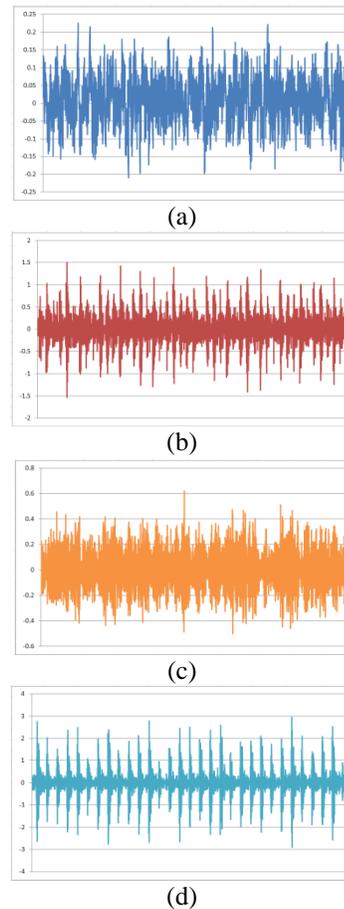


Fig. 3. Vibration signals acquired from four states of bearing. (a) Normal bearings (NB), (b) Outer race fault (ORF), (c) Inner race fault (IRF), (d) Ball fault (BF)

Single point faults with diameter 0.007 inches were introduced separately at inner race, ball, and outer race of the drive end and bearings, using electro-discharge machining. The motor speed and load is 1797 rpm and 2 hp respectively. The bearing vibration signals for normal condition and three fault conditions were collected using a 16 DAT recorder with a sampling frequency of 12,000 Hz and were post-processed in a matlab environment.

Each vibration signal condition was cut into 30 samples with the length 3,600 points. Then, the first 20 samples of each vibration signal condition were used to training and modeling, and the rest were kept for testing the ANNs. Also raw signals of each defect for one sample (3,600 points) are shown in fig. 3.

#### IV. SIGNAL PROCESSING TASK

##### A. Continuous Wavelet Transform (CWT)

The wavelet is obtained from a signal function by  $\psi_{(a,b)}$  translation and dilation

$$\psi_{(a,b)}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2)$$

Where  $a$  is the so-called scaling parameters,  $b$  is the time localization parameter and  $\psi(t)$  is called the "mother wavelet". The parameters of translation  $b \in R$  and dilation  $a > 0$ , may be continuous or discrete. The wavelet transform of a finite energy signal  $x(t)$  with the analyzing wavelet  $\psi(t)$  is the convolution of with a scaled and conjugated wavelet

$$CWT(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi^*\left(\frac{t-b}{a}\right) \quad (3)$$

where  $\psi^*(t)$  stands for the complex conjugation of  $\psi(t)$ . The wavelet transform  $CWT(a,b)$  can be considered as functions of translation  $b$  with each scale  $a$ . Eq. (3) indicates that the wavelet analysis is a time-frequency analysis, or a time-scaled analysis.

##### B. Wavelet Selection

In this stage selecting an appropriate mother wavelet to extracting feature based CWT coefficients using three basis criteria considered.

1) *Maximum Wavelet Energy criterion (MWEC)*: Time-Scale domain signal energy shows the similarity between signal and wavelet which be selected. An appropriate wavelet selected as the base wavelet, must have maximum amount of energy of the wavelet coefficients [12], [13]. The energy at each resolution level  $n$  is given by

$$E(n) = \sum_{i=1}^m |C_{n,i}|^2 \quad (4)$$

where 'm' is number of wavelet coefficients and  $C_{n,i}$  is the  $i$ th wavelet coefficient of  $n$ th scale.

The total energy can be obtained by

$$E_{total} = \sum_n \sum_{i=1}^m |C_{n,i}|^2 = \sum_n E(n) \quad (5)$$

The total wavelet energy of each sample which depends to particular class was obtained, and then mean of samples was assign as a total energy which given by

$$E_{ts} = \frac{1}{N} \sum_j E_{total} \quad (6)$$

where 'N' is number of samples.

2) *Minimum Shannon Entropy criterion (MSEC)*: An appropriate wavelet is selected as the base wavelet, which minimizing the Shannon entropy of the corresponding wavelet coefficients [14]. Shannon entropy of signal is given by

$$S_{entropy}(n) = -\sum_{i=1}^m p_i \log_z p_i \quad (7)$$

3) *Maximum Energy to Shannon Entropy ratio Criterion (RC)*: An appropriate wavelet is selected as the base wavelet, which can extract the maximum amount of Energy while minimizing the Shannon entropy of the corresponding wavelet coefficients [14]. A combination of the Energy and Shannon entropy content of a signal's wavelet coefficients is denoted by Energy to Shannon Entropy ratio and is given as

$$\xi(n) = \frac{E(n)}{S_{entropy}(n)} \quad (8)$$

##### C. Feature Extraction Based CWC

In intelligent systems, feature extraction is of paramount importance. Since there is no clear-cut rule for practical vibration signals, a reliable feature is the main factor. We applied continuous wavelet transform (CWT) in this research. The basic theory of CWT has been discussed in stage A. Continuous wavelet coefficients (CWC) show how well a wavelet function correlates with the signal, supposing signal energy and wavelet function energy are equal to one. In this research, CWC was calculated for bearing segmented signals with 87 mother wavelet candidates from different wavelet families: Haar, Daubechies, Symlet, Coiflet, Gaussian, Morlet, complex Morlet, Mexican hat, bio-orthogonal, reverse bio-orthogonal, Meyer, discrete approximation of Meyer, complex Gaussian, Shannon, and frequency B-spline wavelets. To select the most similar mother wavelet for each setup, 4×30 sample signals were processed for faulty and normal conditions. Three wavelet selection criteria which discussed in this section are used and compared to select an appropriate wavelet for feature extraction.

Based on three wavelet selection criteria, bio-orthogonal3.1 wavelet, complex Morlet1-1 wavelet and bio-

orthogonal3.3 wavelet are selected as the best base wavelet among the other wavelets considered. Results are shown in fig. 4. The CWC of all the 120 signals with these wavelets as a base wavelet are calculated at seventh level of decomposition ( $2^7$  scales). When apply the wavelet transform to a signal, if a major frequency component corresponding to a particular scale exists in the signal, the corresponding wavelet coefficients at that scale will have relatively high magnitudes. The statistical features of the continuous wavelet coefficients

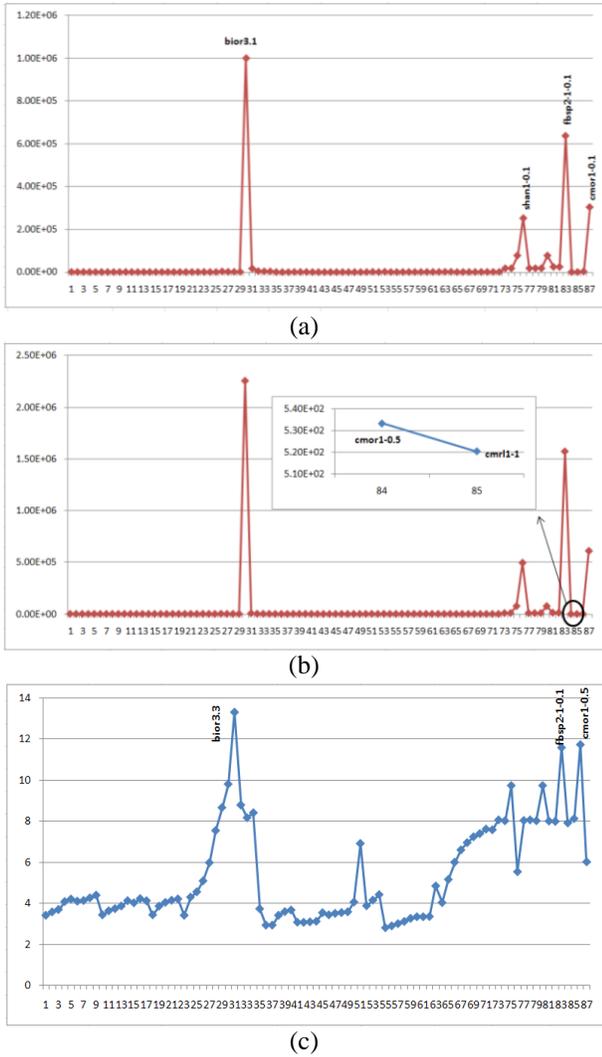


Fig. 4. Wavelet Selection Criteria. (a)The mother wavelets vs. Maximum Energy, (b) The mother wavelets vs. Minimum Shannon Entropy, (c) The mother wavelets vs. Minimum Shannon Entropy.

(CWC) corresponding to the selected scale is calculated. Root mean square (RMS) value, crest factor, kurtosis, skewness, standard deviation, etc., are most commonly used statistical measures used for fault diagnosis of rolling element bearings. Statistical moments like kurtosis, skewness and standard deviation are descriptors of the shape of the amplitude distribution of vibration data collected from a bearing, and have some advantages over traditional time and frequency analysis, such as its lower sensitivity to the variations of load and speed, the analysis of the condition

monitoring results is easy and convenient, and no precious history of the bearing life is required for assessing the bearing condition. When selecting certain normalized statistical moments to monitor the bearing condition, we usually need to consider two most essential characteristics, i.e. sensitivity and robustness. By rectifying the signal, Honarvar and Martin compared the third moment, skewness, of the rectified data to kurtosis, and found that this third moment has better characteristics than kurtosis [15]. In present paper, author's use statistical moments like kurtosis, skewness and standard deviation as features to effectively indicate early faults occurring in rolling element bearing. These statistical features are briefly described as follows:

**Kurtosis:** A statistical measure used to describe the distribution of observed data around the mean. Kurtosis is defined as the degree to which a statistical frequency curve is peaked.

$$Kurtosis = \left\{ \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum \left( \frac{x_i - \bar{x}}{s} \right)^4 \right\} - \frac{3(n-1)^2}{(n-2)(n-3)} \quad (9)$$

**Skewness:** Skewness characterizes the degree of asymmetry of distribution around its mean. Skewness can be negative or positive.

$$Skewness = \frac{n}{(n-1)(n-2)} \sum \left( \frac{x_i - \bar{x}}{s} \right)^3 \quad (10)$$

**Standard deviation:** Standard deviation is measure of energy content in the vibration signal

$$Std = \sqrt{\frac{n \sum x^2 - (\sum x)^2}{n(n-1)}} \quad (11)$$

Also a maximum wavelet coefficient (MWC) of each sample is considered as another feature for four different conditions. These statistical features are fed as input to the ANN for faults classification.

TABLE I  
Effects of input signals identification of machine condition with four features based on Wavelet Criteria using ANN (sets: MWC(1), Kurtosis(2), Skewness(3), Std(4))

Case no.	Feature set	MEC (BIOR3.1)	MSEC (CMORL1-1)	RC (BIOR3.3)
1	1	64.8%	62.3%	70.4%
2	2	80.3%	72.3%	75.6%
3	3	93.5%	76.5%	86.1%
4	4	94.5%	85.4%	81.65%
5	1,2	85.45%	78.15%	68.3%
6	1,3	87.5%	70.3%	60.2%
7	1,4	76.3%	68.8%	78.2%
8	2,3	68.5%	89.5%	86.7%
9	2,4	70.1%	91.1%	76.4%
10	3,4	83.7%	90.7%	79.7%
11	1,2,3	75.00%	69.8%	73.3%
12	1,2,4	78.2%	92.2%	83.2%
13	1,3,4	83.5%	83.5%	96.5%
14	2,3,4	99.1%	84.3%	87.3%
15	1,2,3,4	98.2%	99.68%	99.2%

## V. RESULTS AND DISCUSSION

The features which described in pervious section extracted from the CWC, and used as inputs to the ANN and the results obtained are shown in Table II. Total 15 set states according to Eq. (12) were considered.

$$\sum_{n=1, F=4}^4 \frac{F!}{(F-n)!n!} = 15 \quad (12)$$

where 'F' is the number of features.

TABLE II  
A compressive study between the present work and some recent publications.

	References				
	Paya et al. [16]	Abbasian et al. [17]	Kankar et al. [8]	Rafee et al. [7]	Peresent work
Objects	Bearings and gears	Rolling element bearings	Rolling element bearings	Bearings and gears	Rolling element bearings
Defects considered	Defects on inner race of bearing and gear tooth irregularity	Bearing looseness, defects in rolling elements and bearing raceways	Spall in inner race, outer race, rolling element and combined component defects	For gear broken-tooth gear, slight-worn gear and medium-worn gear. For bearing defects, inner race fault, ball fault, and cage fault	inner race fault, ball fault, and outer race fault
Techniques used for vibration signature analysis	D4 wavelet	Meyer wavelet	Meyer, Coiflet5, Symlet2, Gaussian, complex Morlet and Shannon wavelets	324 mother wavelet, Haar, Daubechies, Symlet, Coiflet, Gaussian, Morlet, complex Morlet, Mexican hat, bio-orthogonal, reverse bio-orthogonal, Meyer, discrete approximation of Meyer, complex Gaussian, Shannon, and frequency B-spline wavelets	87 mother wavelet, Haar, Daubechies, Symlet, Coiflet, Gaussian, Morlet, complex Morlet, Mexican hat, bio-orthogonal, reverse bio-orthogonal, Meyer, discrete approximation of Meyer, complex Gaussian, Shannon, and frequency B-spline wavelets
Features considered	10 wavelet numbers indicating both time and frequency and their 10 corresponding amplitudes	Fundamental cage frequency (Fc), ball pass inner raceway frequency (FBPI), ball pass outer raceway frequency (FBPO) and ball rotational frequency (FB)	Statistical features namely, kurtosis, skewness and standard deviation from wavelet coefficients corresponding to scale maximizing Energy to Shannon entropy ratio or Relative Wavelet Energy	Statistical features namely, kurtosis, skewness, standard deviation, variance, kurtosis, and fourth central moment of continuous wavelet coefficients of synchronized vibration signals (CWC-SVS).	Statistical features namely, kurtosis, skewness, standard deviation, and max wavelet coefficient from wavelet coefficients corresponding to scale maximizing Energy to Shannon entropy Ratio, Max Energy and Min Entropy
Classifier used	Artificial neural networks	Support vector machine	Support vector machines, artificial neural networks, self-organizing maps	Artificial neural networks	Back propagation neural network
Classifier efficiencies	96%	100%	Best efficiency obtained using test set with Meyer wavelet and SVM – 98.6667%	NA	Best efficiency obtained using test set with Complex morlet wavelet and BP – 99.68%

The training success for each was 100%. The results are shown that, case no.15 is appropriate feature set for fault diagnosis. Also based on wavelet selection criteria MSEC and RC are more effectiveness than MEC and MSEC is

appropriate for wavelet selection and its feed back in Fault diagnosis according to results is dominants.

Table I shows accuracy associated with each criterion for faults classification. The correctly classified using MSEC and RC criteria for ANN are 99.68% and 99.2 respectively. These criteria which mentioned are well as a criterion for wavelet selection but MSEC criterion is more appropriate and efficient.

To show the efficiency of the selected features and the methodology, a comparison between the current work and some published literatures has been shown in Table II.

TABLE III  
Wavelet selection based criteria

Criterion	Magnitude	Best Wavelet
Maximum Energy(MEC)	1e+6	Bior3.1
Minimum Shannon Entropy(MSEC)	2.05e-15	Cmor1-1
Maximum Energy to Shannon Entropy ratio(RC)	13.32	Bior3.3

## VI. CONCLUSION

This study presents a methodology for detection of bearing faults by classifying them using back propagation (BP) ANN this methodology incorporates most appropriate features, which are extracted from wavelet coefficients of raw vibration signals. Three wavelet selection criteria Maximum Energy, Minimum Shannon Entropy, and Maximum Energy to Shannon Entropy ratio are used and compared to select an appropriate wavelet for feature extraction. Results obtained from the three criteria for 87 different mother wavelets are shown in Table I and Table III and concluded that the wavelet selected using Minimum Shannon Entropy ratio criterion (Complex Morlet 1-1 wavelet) gives better classification efficiency. The results show the potential application of proposed methodology with machine learning techniques for the development of on-line fault diagnosis systems for machine condition.

## REFERENCES

- [1] T. Kaewkongka, Y. H. J. Au, R. T. Rakowski, and B. E. Jones, "A comparative study of short time Fourier transform and continuous wavelet transform for bearing condition monitoring," *Int. J. COMADEM*. 6 41–8, (2003).
- [2] B.S. Kim, S. H. Lee, M. G. Lee, J. Ni, J. Y. Song, and C.W. Lee, "A comparative study on damage detection in speed-up and coast-down process of grinding spindle-typed rotor-bearing system," *J. Mater. Process. Technol.* 187 30–6, (2007).
- [3] W. Wang, F. Ismail F, and F. Golnaraghi, "Assessment of gear damage monitoring techniques using vibration measurements," *ech. Syst. Signal Process.* 15 905–22,(2001).
- [4] Q. Hua, Z. Hea, Z. Zhanga, and Y. Zia, "Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble," *Mechanical Systems and Signal Processing*. 21 (2007) 688–705.
- [5] A. Widodo, E.Y. Kim, J.D. Son, B.S. Yang, A. C. C. Tan, D.S. Gu, B.K. Choi, J. Mathew, "Fault diagnosis of low speed bearing based on relevance vector machine and support vector machine," *Expert Systems with Applications*. 36 (2009) 7252–7261.
- [6] B. Samanta, and K. R. Al-Balushi, "Artificial neural network based fault diagnostics of rolling element bearings using time-domain

- features,” *Mechanical Systems and Signal Processing* (2003) 17(2), 317–328.
- [7] J. Rafiee, M.A. Rafiee, and P.W. Tse, “Application of mother wavelet functions for automatic gear and bearing fault diagnosis,” *Expert Systems with Applications* 37 (2010) 4568–4579
- [8] P.K. Kankar, Satish C. Sharma, and S.P. Harsha, “Fault diagnosis of ball bearings using continuous wavelet transform,” *Applied Soft Computing* 11 (2011) 2300–2312
- [9] I. Daubechies, “Ten lectures on wavelets,” *CBMS-NSF Series in Applied* (1991).
- [10] S. Haykin, “*Neural Networks. A Comprehensive Foundation*,” Pearson Prentice Hall Publications, Ontario, Canada, 2005.
- [11] Case Western Reserve University Bearing Test Data Center. <http://www.eecs.case.edu/laboratory/bearing/>
- [12] O.A. Rosso, and A. Figliola, “Order/disorder in brain electrical activity,” *Revista Mexicana De Fisica* 50 (2004) 149–155.
- [13] O.A. Rosso, S. Blanco, J. Yordanova, V. Kolev, A. Figliola, M. Schurmann, and E. Basar, “Wavelet entropy: a new tool for analysis of short duration brain electrical signals,” *J urnal of Neuroscience Methods* 105 (2001) 65–75.
- [14] R. Yan, “Base wavelet selection criteria for non-stationary vibration analysis in bearing health diagnosis,” *Electronic Doctoral Dissertations for UMass Amherst, Paper AAI3275786*, <http://scholarworks.umass.edu/dissertations/AAI3275786>, January 1, 2007.
- [15] F. Honarvar, and H.R. Martin, “New statistical moments for diagnosis of rolling element bearing,” *Journal of Manufacturing Science Engineering* 119 (1997) 425–432.
- [16] B.A. Paya, I.I. Esat, M.N.M. Badi, “Artificial neural networks based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor,” *Mechanical Systems and Signal Processing* 11 (1997) 751–765.
- [17] S. Abbasion, A. Rafsanjani, A. Farshidianfar, N. Irani, “Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine,” *Mechanical Systems and Signal Processing* 21 (2007) 2933–2945.