International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 14
October2016

Bearing Fault Diagnosis Based On W P T And Automatic Reconstruct The Raw Signal With FF And BRA Of Artificial Neural Networks

Anas Hamidk. Aljemely, Jianping Xuan, Salman Khayoon, Qamar Ud D. Abid
School of Mechanical Science and Engineering, Huazhong University of Science and Technology, China
E-mail: anasaljemely@yahoo.com

Abstract: *The intricate signal comprised of a machine component, that has several sources of vibration generation, leads to further investigation in signal processing and fault classification of ball bearing to improve maintenance systems efficiency. In this paper, we present an approach to extract time-domain features by wavelet packet WPT technique that is used to obtain the restored signal automatically from the original signal analyzed previously relying on robust energy changing of the WPT's coefficients. The training and testing phases are conducted by fed the features as samples into Bayesian Regularization Algorithm of neural networks BRANNs. The effectiveness of proposed method is evaluated by comparison with FFNNs. The observation results from the experimental data proved that the time feature and BRANNs are very effective for classifying the ball bearing conditions as IRF, ORF, BSF and healthy bearing. Consequently, this approach has the capability to realize categorization for the signals of rotating machines in fault diagnosis systems.*

Keywords:WPT decomposition, automatic sub-band selection, time features extraction, FFNNs and BRNNs algorithms, bearing conditions classification.

1. Introduction

Rolling element bearings are used as an important element in machines and used primarily to support rotating shafts in mechanical equipment and their failure one of the most frequent reasons for machine breakdown. However, the vibration induced signals that generated by faults in defective bearing have been widely studied, and very powerful diagnostic techniques are now available to discuss[1]. The personal damage and financial loss perhaps occurs because of an element of the bearing can refracted if the fault is not detected and diagnosed well in advance. Proper operation in the applications depends on, to a great life extent, the smooth and quiet running of the bearings [2]. The vibrations released by rotating machines are probably to express an interaction between time- and angle-dependent components. For instance, thesignal can be viewed as a series of cyclic impacts locked to the shaft angle and exciting structural resonances [3] Fault diagnosis of rolling element bearings using signal processing of vibration is widely used to prevent incidentsin machinery. In the vibration signals decomposition field and best solution obtain, several techniques have been investigated in different domains including time domain, frequency domain and time-frequency domain [4].The most commonly exploited methods in time-domain of vibration signal are interpreted over several parameters. Statistical parameters are employed to depict the overall information of the signal and it can be taken out from vibration signal such as RMS, Crest factor, peak value, mean value, probability density function, statistical moments and others [5]. Whereas, in a

International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 14
October2016

frequency domain using Fourier transformations analysis function that is working to transform time domain signals into the frequency domain. The key point in time domain and frequency analyses is that the direct use of informational content in one domain is excluded when the other domain is employed [6].Most of common techniques relied on the hypothesis of constant – or possibly variation but in a stationary way of operationspeed [7]. However, when initial bearing defects occur, the vibration signals generated often display non-stationary and non-linear behaviors. Therefore, in this paper, wavelet packet transforms WPT and artificial neural networks ANNs have been developed and applied to the fault diagnostic of a bearing. The basic theory of wavelets is decomposition the signalinto various sub-frequency bands and the selection of a suitable mother wavelet for that purpose [8]. WPT with Daubechies mother wavelet is applied on the signalsare measured from experimental bearing conditions (normal and faulty).WPT is decomposed the raw signals into several levels over many sub-frequency bands are then automatically selected based only upon the highest energy of these sub- bands to reconstruct the restored signal. This approach is conducted to exclude the un-useful information from the original signal and a filtered signal with useful information is obtained. A features vector is estimated by using the time domain statistical parameters of the restored signals. These features show that a training of neural network can diagnose rolling bearing fault conditions. Poor generalization capability or over-fitting problem happens when a neural network over learns during the training period. Consequently, in spite of being well trained, the network may not perform well on unrevealed data set because of its deficiency in the capability of generalization. We proposed Neural Network with Bayesian Regularization algorithm to overcome these problems and prove its ability in the field of nonlinear pattern classification for element bearing conditions. This approach is based on selected segments of available experimental data. Our proposed method is developed and implemented in MATLAB program.

2. Wavelet Packet Transform (WPT)

Wavelet is a convolution technique with segments of a known signal to extract information from the unknown signal. The wavelet packet transform is more suitable than other wavelets in the signal analysis fields because of its property to identify and classify the frequency domain. It is one of the expert tools for signal analysis because it possesses abilities to achieve higher discrimination by decomposing the higher frequency domain [9]. It can analyze and reconstruct a signal rigorously. Compared with wavelet analysis, WPT can not only divide the low-frequency part, but it is also able to realize a secondary decomposition of the high-frequency portion and therefore, it is powerful in the signal analysis[10].The WPT is a generalization of the wavelet transform by defining two functions as $W_0(t) = \phi(t), W_1(t) = \psi(t)$ where $\phi(t)$ and $\psi(t)$ are the scaling and wavelet functions respectively. In case of orthogonal, the function can be described as

$W_n(t)$,n=0,1,2,…..and T =0,1,2,…..

$$W_{2n}^{j} = \sqrt{2}\sum_{k} h(T) W_n^{j}(2t - T)$$

(a.1)

International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 14
October2016

$$W^{j}_{2n+1} = \sqrt{2}\sum_{k} g(T) W^{j}_{n}(2t - T)$$
(b.1)

$$W_{j,n,T}(t) = 2^{-j/2} W_{n}(2^{-j}t - T)$$
(1)

Where j is a scale parameter, T is a time localization parameter, $W_{j,n,T}$ represents the analysis function of wavelet packet. n levels of Wavelet Packet Transform have decomposed the original signal into 2^n bandwidth that is equal to the frequency bands in a result. At the same time, the varying of energy is considered as a basis for selecting reconstruction frequency band to restore the original signal again. This process can improve the accuracy because it excludes the unwanted residual signals having no useful information in fault diagnosis method. In this work WPT is applied on measured signal by implementing 3 levels and select a robust energy variation of frequency bands through 8 sub-bands of the third layer. We can describe the wavelet analyzing and restoring process by two steps. First: Determining the reconstruction signal $X_{R3,j}$ of each frequency band for the original signal decomposed for three levels and then calculating the energy and its distribution of each reconstructed signal, defined as:

$$E_{3,j} = \int \left| X_{R3,j}(t) \right|^2$$
(2)

Second: calculate the restored signal $S_{res}$by selecting frequency bands which possess robust variation and highest amount of energy. Thus, we can get the restored signal by gathering only these frequency bands together. This process is described as:

$$S_{res} = \sum X_{R3,j}$$
(3)

Where j =0,1,2,..., $2^3 - 1$ is number of coefficient at each node in 3 levels of WPT decomposition.

3. Artificial Neural Networks (ANNs)

Artificial Neural Network is a data computing processing technique that is designed to imitate the human brain and nervous systems. An ANN consists of large number of interconnected artificial Processing neurons called nodes which are configured to solve a specific problem in different applications, each node is connected to others by weight element called synapses and organized into layers to formation of the network. The ANN consists of main layers which are, an input layer, a hidden layer, and an output layer. The number of formatted nodes in each input and output layers are constructed by the nature of the problem that is needed to be solved and relying on the number of input and output variables. A trial and error process is done to specify the nodes and hidden layer numbers dependent upon good results in training phase.

Each neuron in every layer of artificial neural network calculates the sum of the input signals $x(t)$, the sum then goes through an activation function (*f*). The following equation defines the output of the network (*O*).

$$O = f\left\{\sum_{i=1}^{n} w_i . x_i(t) + b\right\}$$
(4)

Where w_i is the corresponding weight value and *b* is a bias vector. The activation function *f* transforms the weighted inputs into the output O .The sigmoid or logistic function is the most commonly used function due to it limits the nonlinear functions and the keeps the data outputs between 0 and 1. Back propagation method is known for its common use in training the network [11]. The back propagation algorithm for an ANN is a generalization of the lower mean square algorithm to adjust the network parameters in order minimizes the mean square error (MSE). Therefore, it is necessary to choose a better algorithm to increase the generalization capability and overcome the over-fitting problem to the data inputs. In this work, multilayer perceptron neural network MLPNNs together with back propagation learning is used first and then followed by Bayesian Regularization Algorithm BRA which is computed to overcome a poor generalization of the data used. It is compared with feed-forward back propagation algorithm neural networks FFNNs to prove the efficiency of BRA modification for the error in training phase of the networks.

3.1. A feed-forwardNeuralNetworks (FFNNs)

The feed forward neural network is considered primarily as the simplest kind of artificial neural network contrived. Basically in this network, the information transfers are only in one path, forward. In class of multi layers percepton networks MLPNNs, it is comprised of multiple layers of computational units, habitually interconnected in a feed-forward way. These networks apply a sigmoid function in its applications as an activation function. MLPNNs have advantage to compute a continuous output instead of a step function. One of the best choices is the so-called logistic function or sigmoid function as: $g(x) =1/1+e^{-x}$ .The benefit of continuous derivative for this function is its easy computation that leads to its use in back-propagation algorithm as: $g' =g\ (1-g)$. For neural networks conditions, the universal approximation theorem that defines: in every continuous function that maps intervals of real numbers to some output interval of real numbers can be approximated arbitrarily by a multi-layer perceptron with just one hidden layer. For example, in the sigmoidal functions, this result embraces for a wide range of activation functions; the most popular of learning techniques is back-propagation that is used in multi-layer networks. To compute the value of the error-function, the correct answer is compared with the output values by using various techniques, the error could feed back through the network. Otherwise, the algorithm adjusts the weights through utilizing this information of each connection in order to reduce the value of the error function by some small amount. After the recurring the process for suitedthe large number of training sequences, thus, the error of the calculations will be small due to the network concourse to some state. In this case, the network maybe considered has learned a certain target function. To adjust weights appropriately, one stratifying way is a general method of optimal non-linear that is called gradient descent. For this purpose, the derivative of the error function with a network weights is computed, and only the back-propagation is applied in networks with differentiable activation functions because the weights are then changed in such a way that the error is reduced. Overall, the problem of training a network to achieve good results is a reasonably subtle issue that requires additional techniques even on samples that were not used as training samples. This is

International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 14
October2016

more significant for cases where only restricted numbers of training samples are obtainable. [12] The over-fits of the network training information are the big peril which fails to apprehend the true statistical process creating the data. Theory of computational learning is fascinated with training classifiers on a limited quantity of information. In the setting of neural networks a simple heuristic, called early stopping, frequently certifies that the network will generalize well to instances that are not available in the training set.

3.2. Bayesian Regularization NeuralNetworks (BRNNs)

Unfortunately, the artificial neural networks system has Over-fitting problem or poor generalization capability, especially during the training period. A neural network while working in learning perhaps does not carry out well on invisible data set due to its lack of generalization capability [13]. We therefore, suggested Bayesian Regularization Algorithm approach to overcome this problem. It decreases the over-fitting problem by enhancing the account by the goodness-of-fit data as well as the network model. BRA approach usually modulates the used objective functions, such as the sum mean squared error of networks (MSE). MSE is symbolized as E_N and defined as:

$$E_N = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 \qquad (5)$$

Our modulation goal is to improve the model's generalization capability. In fact, we can extend the objective function by sum of squares of the network weights, and E_w becomes

$$F = \beta E_N + \alpha E_w \qquad (6)$$

Where α and β are the assumed weights and biases, respectively, of the network which is to be optimized in Bayesian framework [14]. This approach is considered to perform the adjustment the error of back-propagation learning which passes through different layers of the network: forward and backward. Selecting the network kind is rely on the problem to be dealt with and solved. Therefore, a back propagation by Bayesian Regularization Algorithm with Levenberg-Marquardt algorithm for neural network training is considered in this study [15].Classification process by BRANNs for bearing conditions are descripted in fig. 1

Input layer

Hidden layer

Classified Bearing Conditions

International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 14
October2016

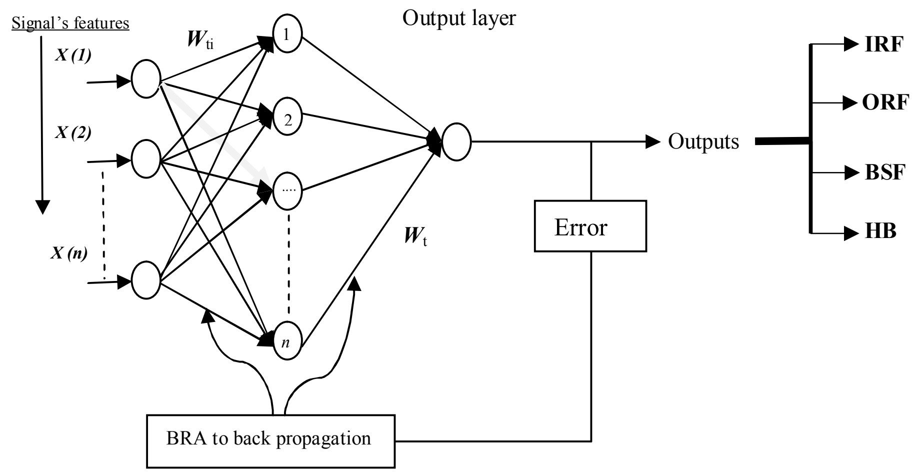


***Figure1**. Schemeof Back propagation of BRA for neural networks classification process.*

3.3. The BRA neural networks proposed

ANN has been very effective for data remediation and information processing and it has high level of adaptability that cannot be obtained from another numerical procedure. Therefore, we utilized this technique to classify the data that is obtained from defective and healthy bearing conditions. As mentioned previously, an ANN has 3 layers in a general input, hidden and output. ANNs with BRA is designed for ten input nodes, five-thirty five nodes for hidden layer and four nodes for the output layer. Thus, this number of nodes is considered for the selected bearing conditions that are to be classified: IRF, ORF, BSF and healthy bearing HB in this application. The training of BRANNs can be ceased according to the neural network settings when reaching our target which can be either mean square error (MSE)of 10^{-7} or maximum iteration number(epoch)of 1000. Assigning incipient weights and biases values of the network are randomly generated by the program, and the targets of output layer nodes are imposed as

Inner race fault [1 0 0 0]
Outer race fault [0 1 0 0]
Ball spin fault [0 0 1 0]
Normal condition [0 0 0 1]

Each value of the fourth conditions is supposed to be the ideal value of output layer nodes which is compared with output values and then adjusted by an ANN proposed until a good result is obtained in close approximation to the equivalent target values.

4. Experiment set –up

For the purpose of discovering and implementing signal analysis theories and techniques, Case Western Reserve University published bearing signal data on-line for researchers [16]. The basic description of the test rig is a 3 HP Electric motor leading a shaft with a mountedtorque transducer (Accelerometer). The dynamometer is driven by the applied Torque of a shaft, while the electronic control speed system and test bearings are supported on the motor shaft. Artificial faults are designed in diameter from 0.007 to 0.028 inches classified in to the 6205-2RS JEM drive-end and fan-end model 6203-2RS JEM. Both bearings used are SKF deep groove ball bearings. The artificial faults are introduced on the elements of the bearing including, the inner raceway, outer raceway and roller element. In test rig operation of each defective bearing is reassembled individually and vibration signals are measured at motor speed domain (1797 to 1720).

5. Application

The accelerometersare mounted on the motor housing with magnetic bases and fixed on the drive end bearing side of the testing set up. In this paper, the data used as vibrational signals, collected by the previously mentioned using test model 6205-2RS JEM deep groove ball bearing. The bearing geometry was: 39.04mm bearing diameter; 7.94mm ball diameter; 9 number of balls and defect frequencies were calculated from that geometry with a 1797rpm rotational speed of the bearing. Table 1 describes the equations which were used to compute the defect frequency for each element. The sampling frequency and fault sizes were 12 kHz, and 0.007 inches respectively. Four conditions are selected, including three fault vibration signals; outer race vibration signal, inner race vibration signal, roller spin vibration signal, and the healthy bearing vibration signal.

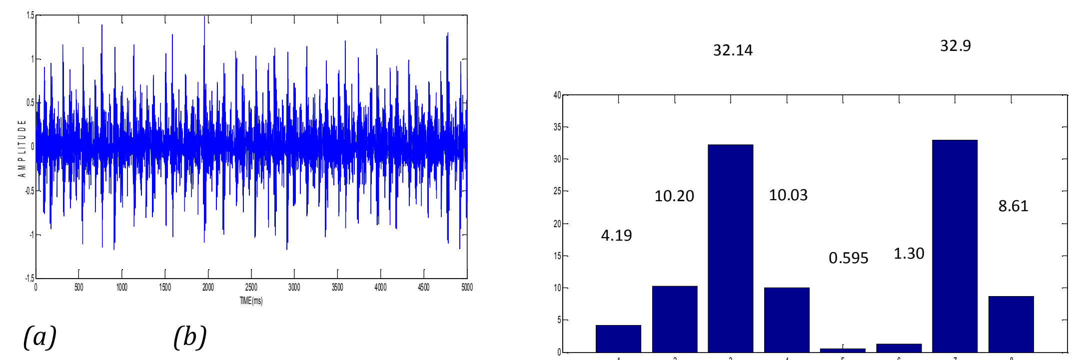
Table1. *Bearing elements defect frequency equations.*

1	Outer race defect frequency(**ORF**)	$f_{BPO} = \frac{n}{2} f_r (1 + \frac{d}{D}\cos\beta)$
2	Inner race defect frequency(**IRF**)	$f_{BPI} = \frac{n}{2} f_r (1 - \frac{d}{D}\cos\beta)$
3	roller spin defect frequency(**BSF**)	$f_{BPR} = \frac{D}{d} f_r \left[1 - (\frac{d}{D}\cos\beta)^2\right]$

Where $fr$is a shaft frequency calculated as rpm/60, n is a number of balls in bearing, D is the outer diameter of roller bearing, d is the inner diameter of the ball bearing and β is a ball contact angle. Thus the equations can calculate, mathematically, the characteristic frequency of each bearing element fault as outer race fault (f_{BPO}), inner race fault (f_{BPI}) and roller frequency fault (f_{BPR}).

5.1. Time-domain features extraction

Analysis of the raw vibration signal in time domain is based on WPT decomposition to calculate the restored signal by eq.3 to combine the corresponding sub frequency bands which are selected automatically based on high energy. Fig.2 shows the original signal and its energy percentage of each frequency band signal of the inner race fault by wavelet packet decomposition. This procedure is proposed in this paper to extract the time-domain features.



(a) *(b)*

Figure2. *(a) Time domain of inner race fault signal, (b) Energy percentage of each sub band of decomposed signal by WPT.*

Substantial change in energy of each sub band, in vibration signal decomposed, is directly related to the defect of bearing element. This change indicates the bearing condition. Extraction of the statistical parameters is considered to diagnose the fault. The features used are: peak value (P_v), mean (M_F), variance (V_f), root mean square (RMS), square mean root (SMR), peak-peak (PP), skewness (SK), kurtosis (K_v), crest factor (CR_f), latitude factor (LA_f). The 10 time-domain features are calculated mathematically in MATLAB. The mathematical calculation of these features is expressed in table 2 as

Table2.*Feature equations of the time domain.*

1	$P_v = [\max x(n)]$	6	$PP = \frac{1}{2}[\max(x(n)) - \min(x(n))]$
2	$M_f = \frac{1}{N}\sum_{n=1}^{N} x(n)$	7	$SK = \frac{\sum_{n=1}^{N}[x(n) - M_f]^3}{(N-1)V_f^{\,3}}$
3	$V_f = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x(n) - M_f)^2}$	8	$K_v = \frac{\sum_{n=1}^{N}[x(n) - M_f]^4}{(N-1)V_f^{\,4}}$
4	$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(x(n))^2}$	9	$Cr_f = \frac{\max\lvert x(n)\rvert}{RMS}$
5	$SMR = \left[\frac{1}{N}\sum_{n=1}^{N}\sqrt{\lvert x(n)\rvert}\right]^2$	10	$LA_f = \frac{\max\lvert x(n)\rvert}{SMR}$

Where $x(n)$ is a signal series (n =1, 2,..., N) and N is number of sampling points.

The features 1, 2 and 4, 6 represent the magnitude and energy of the time-domain of vibration signals. 3, 7 and 10 describes the vibration signals spreading in the time-domain [17]. Thus, features can embed the most important information contained in the signal. The total 10-time domain features are proposed as data input vector to the training and testing neural network. The feature values of the first segment for the original signals are shown in table 2.

Table3. *1st segmentinput vector values of time-domain features at each bearing condition.*

Sr. No.	Bearing Condition	Feature extracted values in each bearing condition P_v	M_f	V_f	RMS	SMR	PP	SK	K_v	Cr_f	LA_f
1	IRF	1.61	0.014	0.080	0.281	0.161	2.82	0.114	5.47	5.71	9.99
2	ORF	3.89	0.031	0.412	0.64	0.285	7.42	0.028	7.94	6.06	13.6
3	BSF	0.61	0.015	0.017	0.13	0.093	1.15	-0.014	3.06	4.67	6.53
4	H	0.32	0.012	0.004	0.067	0.05	0.64	-0.056	3.28	4.76	6.43

5.2. Data normalization

In this work, the extracted time domain features of the segments signal data is normalized into a range between0.1 to 0.9, before fed to the neural network according to the following equation

International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 14
October2016

$$X_i = 0.8\left\{\frac{x_i - \min(x)}{\max(x) - \min(x)}\right\} + 0.1 \qquad (7)$$

Where X_i is normalized data, x_i is primary data, $\max(x)$ and $\min(x)$ are the maximum and minimum values of actual data, respectively. We normalized the data input to prevent a discontinuation of the training or slowing down of data processing during implementation of sigmoid function in the training phase; this preparation leads to obtain the further numerical immovability of data processing and improvement in the performance of a neural network.

6. Results and discussion

For the network's training and testing to perform well, the used data of vibration signal for each bearing condition is divided into 4 segments in order to increase the number of samples in each class, these signals are employed to extract ten-time domain parameters. Thus, sixteen signals are obtained for four bearing conditions, and hence the total number of 160 samples was considered in the input matrix of fault classification. Among the samples, 70% were used for training purpose, 15% for validation and 15% for testing purpose. The data were tested after training process to assess the proposed ANNs. In back-propagation process, the most commonly used is FFNNs algorithm. The comparison of BRANNs with FFNNs for performance results has proved and obtained the evaluation of the proposed method. Table 3 presents the performance of BRANNs and FFNNs in different networks with different nodes in hidden layer.

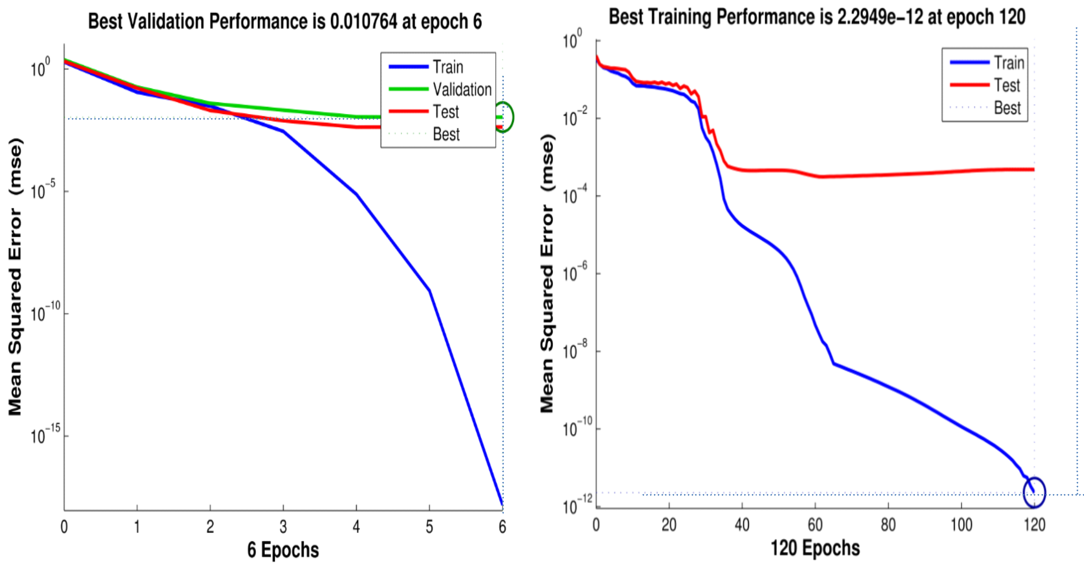
Table3.*(MSE) performance at different number of nodes in hidden layer.*

network	No. of nodes in hidden layer	Performance (MSE)	
		FFNNs	BRANNs
1	5	0.0157	$5.99*10^{-5}$
2	10	0.0165	$1.03*10^{-4}$
3	15	0.0398	$1.43*10^{-4}$
4	20	0.0202	$6.22*10^{-5}$
5	25	0.0082	$1.55*10^{-4}$
6	30	0.0032	$4.89*10^{-4}$
7	35	0.0019	$1.44*10^{-4}$

The value of MSE represents the performance of both algorithms. Proposed BRANNs and FFNNs are trained by different networks with different nodes in hidden layer. Fig.3 presents the performance comparison of the both algorithms from the results, it can be concluded exclusively that both algorithms performed well to decrement the error function in neural networks training. But it is obvious that BRANNs is more effective to treat the problem of errors generated in back-propagation learning until it reaches a minimal amount of generalization.Lowes value of MSE, 0.0019, presents the best performance for FFNNs obtained in network 7 with 35 nodes in hidden layer while for BRANNs it is $5.99*10^{-5}$ in network1 with 5 nodes in hidden layer. Moreover, it is the correlation factor R that enables us to evaluate the

International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 14
October2016

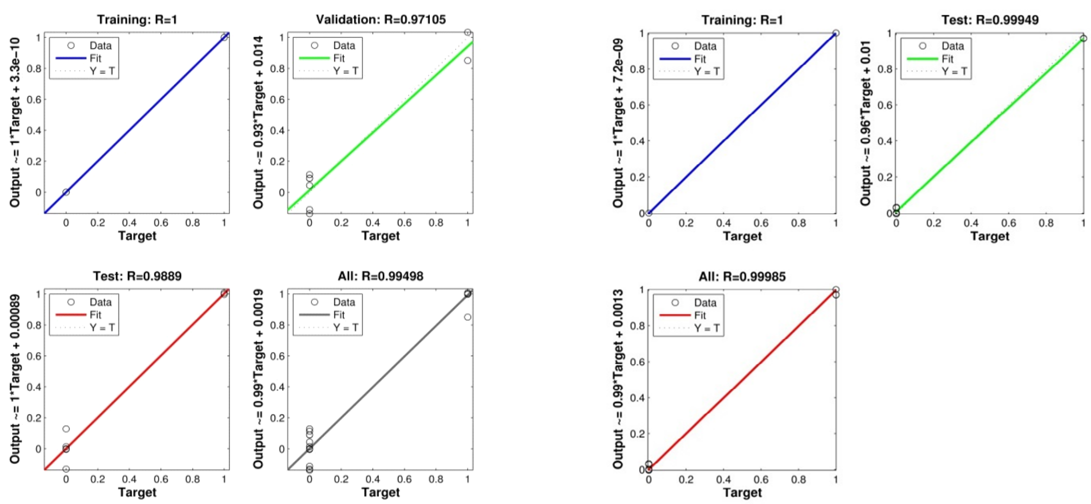
accuracy of each network. Fig.4 depicted the correlation coefficients of FFNN and BRANN algorithms for the data imposed.



(a)(b)

Figure3*. Performance of****a)****FFNNs Training.* ***b)****BRANNs Training.*

.



(a) ***(b)***

Figure4.a) *Correlation coefficients of FFNN performance,* ***b)*** *Correlation coefficients of BRNN performance.*

It is clear that the values of MSE for Bayesian regularization NN are lower than feed forward NN. Thus, BRANNs has performed very well to classify the bearing conditions. The classification

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 14
October2016

accuracy of results for 160 samples has reached 100%.Table 4 enlists the output results of testing phase in accordance with the target specified for each bearing condition.

Table4. *Testing output results of a better BRANN.*

Bearing conditions	target	actual outputs data				classified result
Inner fault signal	1 0 0 0	0.9922	-0.001	0.0085	$4.2*10^{-4}$	IRF
Outer fault signal	0 1 0 0	$2.7*10^{-9}$	1.000	$-1.5*10^{-11}$	$6.9*10^{-8}$	ORF
Ball fault signal	0 0 1 0	$-1.9*10^{-9}$	$-5.8*10^{-7}$	1.000	$-1.7*10^{-8}$	RSF
Healthy bearing	0 0 0 1	$6.2*10^{-5}$	$-1*10^{-6}$	0.0073	0.9927	HB

WPT decomposition and time-domain features were effective in operation of the neural network.The accuracy of the result data of these parameters which are determined corresponding to ideal outputs of BRANN proposed system are demonstrating the capability of this algorithm to classify the roller bearing conditions. It is indicated that using Bayesian Regularization to adjust the error is more suitable than other algorithms through NN processing.

7. Conclusion

In this paper, we investigated an approach for categorization of roller bearing conditions based on an automatic selecting frequency band of a raw signal that is decomposed by wavelet packet transform to extract time domain parameters from restored signal which are applied as input vectors into feed ward back propagation NN, Bayesian regularization of NN to account the goodness-of-fit data through training process of network. BRANNs were very effective overcoming the Over-fitting problem in comparison with FFNNs. This approach was successful for classification all bearing cases relying on the experimental results that reached to 100% of accuracy. It shows that the proposed method possessed a competence to extract the features, classifying ball bearing cases and improve the duty of fault diagnosis systems and make it more effective than FFNNs to reduce the MSE. The future work will be examining this method by applying to other rotating machines.

Acknowledgement

The author is grateful to the support by school of mechanical science and engineering, Huazhong University of Science and Technology, China for providing the research area.

References

[1] Randall, Robert Bond. *Vibration-based condition monitoring: industrial, aerospace and automotive applications*. John Wiley & Sons, 2011.

[2] Tandon, N., and A. Choudhury. "A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings." *Tribology international* 32, no. 8 (1999): 469-480.

[3]Urbanek, Jacek, Jerome Antoni, and Tomasz Barszcz. "Detection of signal component modulations using modulation intensity distribution." *Mechanical Systems and Signal Processing* 28 (2012): 399-413.

[4]Pandya, D. H., S. Upadhyay, and S. P. Harsha. "Ann based fault diagnosis of rolling element bearing using time-frequency domain feature." *International Journal of Engineering Science and Technology* 4, no. 6 (2012): 2878-2886.

[5]Lin, Jing, and Liangsheng Qu. "Feature extraction based on Morlet wavelet and itsapplication for mechanical fault diagnosis." *Journal of sound and vibration* 234, no. 1 (2000): 135-148..

[6]Li, Bo, Gregory Goddu, and Mo-Yuen Chow. "Detection of common motor bearing faults using frequency-domain vibration signals and a neural network based approach." In *American Control Conference, 1998. Proceedings of the 1998*, vol. 4, pp. 2032-2036. IEEE, 1998

.[7]Ocak, Hasan, Kenneth A. Loparo, and Fred M. Discenzo. "Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics." *Journal of sound and vibration* 302, no. 4 (2007): 951-961.

[8]Abboud, D., J. Antoni, S. Sieg-Zieba, and M. Eltabach. "Envelope analysis of rotating machine vibrations in variable speed conditions: A comprehensive treatment." *Mechanical Systems and Signal Processing* 84 (2017): 200-226.

[9]Griffiths, K. R., B. J. Hicks, P. S. Keogh, and D. Shires."Wavelet analysis to decompose a vibration simulation signal to improve pre-distribution testing of packaging." *Mechanical Systems and Signal Processing* 76 (2016): 780-795.

[10]Tang, G. and W. Cai, 2009. Application of wavelet pocket and envelope analysis to fault diagnosis of rolling bearings.Vib. Measurement Diagnosis, 2: 201-204.

[11]Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors." *Cognitive modeling* 5, no. 3 (1988): 1.

[12]Balabin, Roman M., Ravilya Z. Safieva, and Ekaterina I. Lomakina. "Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction."*Chemometrics and intelligent laboratory systems* 88, no. 2 (2007): 183-188.