

Extreme Learning Machine Based on Stationary Wavelet Singular Values for Bearing Failure Diagnosis

Nibaldo RODRIGUEZ¹, Carolina LAGOS^{1*}, Enrique CABRERA², Lucio CAÑETE³

¹ Pontificia Universidad Católica de Valparaíso

Av Brasil 2241, Valparaíso, 2362807, Chile
nibaldo.rodriguez@pucv.cl; carolina.lagos.c@mail.pucv.cl

(*Corresponding author)

² CIMFAV, Universidad de Valparaíso

General Cruz 559, Valparaíso, 2363326, Chile
enrique.cabrera@uv.cl

³ Universidad de Santiago de Chile

Av. Ecuador 3769, Santiago, 9170124, Chile
lucio.canete@usach.cl

Abstract: Intelligence condition monitoring based on vibration signal analysis plays a key role in improving rolling bearings failure diagnosis in mechanical transmission systems. Unexpected failures in the bearings may cause machine breakdowns that are very expensive for the industry. Hence, this study proposes a method to the rolling element bearing failure diagnosis which is based on an extreme learning machine (ELM) algorithm combined with stationary wavelet transform (SWT) and singular value decomposition (SVD). The SWT is used to separate the vibration signals into a series of wavelet component signals. Then, the obtained wavelet components matrix is decomposed by means of a SVD method to obtain a set of wavelet singular values. Finally, the wavelet singular values are used as input to the extreme learning machine for classification among ten different bearing failure types. Obtained results using the proposed model shown high accuracy of diagnosis under variable speed condition.

Keywords: Extreme learning machine, Wavelet analysis, Singular value decomposition, Bearing failure diagnosis.

1. Introduction

Diagnosis of rolling bearings based on vibration analysis is a key element in mechanical transmission systems. Further, early detection of failures on bearings is very important as it improves both reliability and safety of mechanical systems that are widely used in the industry. Thus, in order to accurately classify bearing failures, we need to identify those features that provide us high-quality information regarding the bearing fault characteristic frequency (FCF). Unfortunately, extracting those features from non-stationary and non-linear vibration signals under time-varying speed conditions is not an easy task. Commonly used techniques for features extraction are not very effective and their diagnoses are not very accurate. During the last decade, several time-frequency analysis methods have been applied to the features extraction problem described above. Among them we can find empirical mode decomposition (EMD) [27] and wavelet transform (WT) [9, 21]. Features found by these techniques are then used as the input for classification models which in turn aims to determine the status of the rolling bearings.

The EMD method can decompose signal into a sum of intrinsic mode functions (IMF) according

to the oscillatory nature of the signal [3]. From IMFs, different features can be extracted. IMF energy features [29], IMF envelope spectrum [30] and IMF energy entropy [25] are among the features that can be extracted from IMFs. These features are passed on to some classification method such as neural networks [29] or support vector machines (SVM) [30]. In particular, authors in [25] use IMFs energy entropy to determine whether there exists a failure or not. In case of failure, a vector of singular values is passed on to a SVM in order to determine the type of the failure. The vector of singular values is obtained by means of singular value decomposition (SVD) of the IMF matrix.

A wavelet neural network (WNN) model combined with ensemble EMD for bearing fault diagnosis is proposed in [28]. Here, the more effective IMFs are selected based on the kurtosis value of each IMF. A subset of ten features from both time-domain and frequency-domain is used as input of the WNN for failure classification. The authors in [26] propose a hybrid model for the bearing failure detection problem. This hybrid model uses permutation entropy (PE) to determine whether there is a failure or not. If a

failure exists, then the PE of a subset of selected IMFs is computed and used as the input of a SVM in order to classify the type of the failure as well as its severity. Moreover, WT can decompose a signal into several independent frequency sub-bands and show features of hidden failures [1, 15, 18, 24]. In [8] authors combine WT and EMD to create a new time-frequency analysis method namely empirical wavelet transform (EWT). A comparative study between EWT and EMD for bearing failure diagnosis using acoustic signals is presented in [10]. In that study, authors create an index based on the kurtosis value to select the more effective IMFs. Results in [10] demonstrate that EWT performs better than EMD in terms of the accuracy of the diagnosis. Further, it is shown that EWT is able to efficiently find the frequency and the harmonic components corresponding to the bearing failure feature frequency.

SVM and feedforward neural network (FNN) are widely used to classify different kind of failures in rotatory machines. However, one drawback of these techniques is that they are quite time consuming. In [4, 6, 31] authors use a method called Extreme Learning Machine (ELM) that aims to improve tuning time in a FNN. Since then, ELM has been used in several studies mainly because its efficiency. For instance, in [23] is combined with LMD and SVD to the diagnosis of bearing failure. Here, singular values obtained from the product function matrix are used as input of ELM. It has also been shown that the models combining LMD-SVD-ELM performs better than EDM-SVD-ELM models [23].

In this article we present a hybrid features extraction method that combines stationary wavelet transform (SWT) and SVD. The identification of the failure type and severity is done during this features extraction process by finding those signatures from the vibration signal that represents each failure type. These extracted features are passed on to an extreme learning machine model (ELM). We call our method, SWT-SVD-ELM. We apply our method on the bearing failure diagnosis under variable speed operation problem. While SWT is able to provide local features in both time-domain and frequency-domain as well as it is able to distinguish sudden changes in the vibration signal, singular values are very stable which leads to a more robust and reliable method. Ten different bearing failures types are considered in our experiments. We

apply our method on a set of well-known bearing vibration signals obtained from [12]. We compare the accuracy of the results obtained by our method to those obtained using two different diagnosis method: the well-known EMD-SVD-ELM model and a recently proposed method that combines Fourier coefficients amplitude (FCA) and auto-encoder ELM (AE-ELM).

Remaining sections of this article are as follows: in Section II we describe the components of the SWT-SVD-ELM algorithm that is proposed in this paper. Section III describes the experimental setup and the features extraction phase. A discussion on the results obtained by the SWT-SVD-ELM, the EMD-SVD-ELM, and the FCA-AE-ELM methods is also carried on this section. Finally, in Section 4 some conclusions are presented and some ideas on the future work are outlined.

2. Methodology

2.1 Stationary Wavelet Transform

Stationary Wavelet Transform (SWT) [14, 16, 17] is a wavelet analysis method. It can be seen as an alternative to discrete wavelet transform (DWT) [7, 13]. SWT and DWT share some similarities. The most important one is that both filters, high-pass and low-pass, are applied at each level on the input signal. At the first level of SWT, an input signal $x = \{x(n) = w_{0,0}(n), n = 1, \dots, N\}$ is convolved with a low-pass filter h_1 defined by a sequence $h_1(n)$ of length r and a high-pass filter g_1 defined by a sequence $g_1(n)$ of length r . Both $w_{1,1}(n)$ and $w_{1,2}(n)$, the approximation coefficient and the detail coefficient respectively, are calculated as follows:

$$\begin{aligned} w_{1,1}(n) &= \sum_{k=0}^{r-1} h_1(k)w_{0,0}(n-k) \\ w_{1,2}(n) &= \sum_{k=0}^{r-1} g_1(k)w_{0,0}(n-k) \end{aligned} \quad (1)$$

Since no sub-sampling is performed, the obtained sub-bands $w_{1,1}(n)$ and $w_{1,2}(n)$ have the same number of elements as the input signal $x(n)$. The general process of the SWT recursively continues for $j = 2, \dots, J$ and is given as follows:

$$\begin{aligned}
 w_{j,1}(n) &= \sum_{k=0}^{r-1} h_j(k)w_{j-1,1}(n-k) \\
 w_{j,2}(n) &= \sum_{k=0}^{r-1} g_j(k)w_{j-1,1}(n-k)
 \end{aligned} \quad (2)$$

Filters h_j and g_j are computed by using an operator called dyadic up-sampling. Using this operator a zero value is inserted between each pair of elements that are adjacent in the filter. Thus, the SWT strategy is then completely defined by the pair of filters (both low-pass and high-pass filters) that is chosen and the number of decomposition steps j . For this paper, a pair of Db2 wavelet filters has been chosen mainly due to its low complexity [7], whereas the decomposition at j -th level is given as follows:

$$j = \log_2 \left(\frac{fs/2}{FCF} \right), \quad (3)$$

where f_s represents the frequency sampling of the vibration signal and FCF denotes the bearing fault characteristic frequency [20].

2.2 Extreme Learning Machine

The output of the ELM is obtained as follows:

$$\begin{aligned}
 t_k(n) &= \sum_{j=1}^L \beta_{k,j} h_j(n) \\
 h_j(n) &= \Phi \left[\sum_{i=1}^J (a_{j,i} x_i(n) + b_j) \right],
 \end{aligned} \quad (4)$$

where L represents the number of hidden nodes, $x_i(n)$ denotes the input vector containing J -features values, $\{a_{j,i}, b_j\}$ are the weights of the hidden layer, $\beta_{k,j}$ are output weights of the output layer and $\phi(\cdot)$ represents the hidden nodes activation functions, which is given as follows:

$$\varphi(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

The Moore-Penrose generalized-inverse method (M-P) [22] is used to estimate the output weights

$\beta_{k,j}$, whereas the weights $\{a_{j,i}, b_j\}$ of the hidden layer are randomly assign. The optimal values of the linear weights of the output layer, for any given representation of the hidden weights, are

obtained using (M-P) method as follows:

$$\beta = H^T \left(\frac{I}{C} + HH^T \right)^{-1} T, \quad (6)$$

where T represents the expected output pattern matrix. The expression $(\cdot)^{-1}$ is the Moore-Penrose generalized inverse of the output matrix of the hidden layer, whereas the regulation parameter is set to $C = 10^8$ for all experiments in this paper. We can summarise the ELM method as the following sequence of steps:

- Determine the number L of nodes in the hidden layer and activation function $\phi(\cdot)$;
- Randomly assign the hidden weights $\{a_{j,i}, b_j\}$;
- Using Equation (5), calculate hidden layer's output matrix;
- Using Equation (6), compute the output weights matrix, $\beta_{k,j}$, of the output layer.

2.3 Bearing Fault Diagnosis Method Based on SWT-SVD-ELM

The fault diagnosis algorithm proposed in this study can be summarised as follows:

Firstly, set the size of the sub-signal and divide the raw vibration signal into sub-signals of N data points. Then, calculates the envelope signal using Hilbert transform, which can be computed as

$$x(n) = \sqrt{x^2(n) + H^2[x(n)]}, \quad (7)$$

with $H[\cdot]$ denotes the Hilbert's transform. Once the Hilbert transform has been computed, decompose the envelope signal into J levels by using SWT. After that, decompose the W wavelet coefficients matrix using the SVD method. The SVD method decomposes the W wavelet matrix into a series of mutually orthogonal, unit-rank, and elementary matrices, whose representation is given as follows [11]:

$$W = \sum_{i=1}^{J+1} s_i u_i v_i^T = USV^T, \quad (8)$$

where $U \in \mathfrak{R}^{(J+1) \times (J+1)}$, $V \in \mathfrak{R}^{N \times N}$, S is the $(J+1) \times N$ diagonal matrix and $\{s_1, s_2, \dots, s_J, s_{J+1}\}$ represent the set of singular values of matrix W . After decomposing matrix W , we need to create $(J+3)$ -dimensional vector of features as follows:

$$z = [s_1, s_2, \dots, s_J, s_{J+1}, En_1, En_2], \quad (9)$$

where En represents the Shannon entropy value and is computed as

$$En_1 = -\frac{1}{\log_2(J+3)} \sum_{j=1}^{J+3} p_j \log_2(p_j)$$

$$p_j = \frac{z_j^2}{\sum_{j=1}^{J+3} z_j^2}, \quad (10)$$

$$En_2 = -\frac{1}{\log_2(N)} \sum_{k=1}^N q_k \log_2(q_k)$$

$$q_k = \frac{x^2(n)}{\sum_{n=1}^N x^2(n)} \quad (11)$$

Finally, once we have computed features vector z , we can perform the bearing fault diagnostic strategy using ELM-classifier with singular value features vector as input.

3. Computational Experiments and Discussion

3.1 Experimental Setup

Experimental raw data used in this paper correspond to vibration signals of deep groove ball bearings of 6205-2RS JEM SKF. It can be obtained from [12]. A 2-horsepower (hp) motor, a torque transducer/encoder and a dynamometer are considered in the experimental setup. During the experiments the bearing holds the motor shaft. An accelerometer mounted on the motor housing (as shown in Figure 1), is used to collect vibration signals.

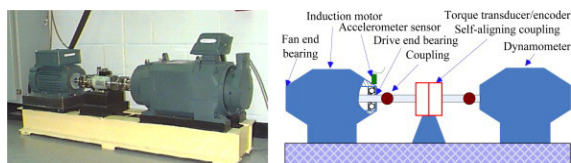


Figure 1. Experimental setup for rolling bearing and its sketch

Single point failures with different failure diameters of 0.007, 0.014, and 0.021 inches are introduced to the driving-end bearing using electro-discharge machining method, with the motor speed varying at 1730, 1750, 1772, and 1797 rpm under four different loads of 0, 1, 2, and 3 horse-power. Digital data is produced at 12,000 samples per second for normal (N) samples and failure samples: inner race fault (IRF), outer race fault (ORF), and ball fault (BF). Further details on the experimental setup can be found in [12].

3.2 Extracting Features for Bearing Failure Diagnosis

The dataset contains 10 bearing health conditions (10 class labels) under rotatory shaft speeds of 1797, 1772, 1750 and 1720 rpm corresponding to loads of 0, 1, 2 and 3 hp, respectively, which are presented in Table 1. The length of the raw vibration signal is set to 120,000 data points (obtained in 10 seconds) and the size of segmentation of sub-signal is set to 1000 data points (83.3 ms), which corresponds to 2.5 times the rotation period. Its period fluctuates between 400 and 418 data points. Therefore, each raw signal is separated into 120 samples and each sample contains 1000 data points. Then, there are 480 samples for each health condition corresponding to variable shaft speed. Afterward, each signal of 1000 data points is decomposed into J levels by using SWT. The J -th value is calculated using the equation (3) and set equal to $J = 3$ since the fault characteristic frequency oscillates approximately between 100 and 200 Hz.

Table 1. Dataset description

# Samples	Fault type	Fault Diameter	Class label
480	N	0	1
480	IRF	0.007	2
480	IRF	0.014	3
480	IRF	0.021	4
480	ORF	0.007	5
480	ORF	0.014	6
480	ORF	0.021	7
480	BF	0.007	8
480	BF	0.014	9
480	BF	0.021	10

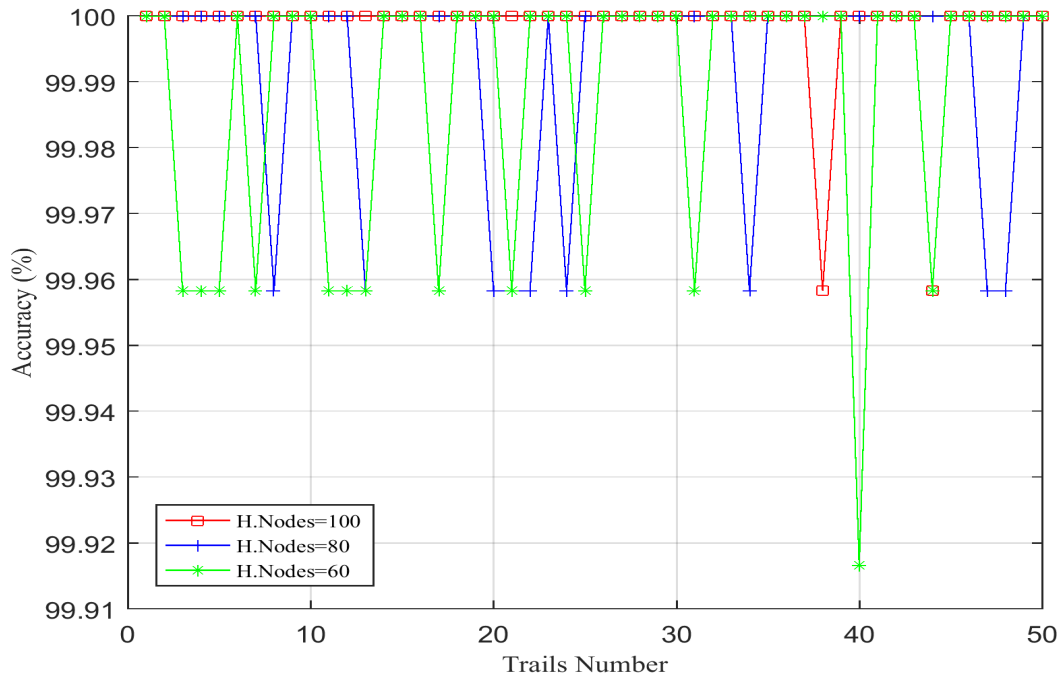


Figure 2. Classification accuracy with SWT-SVD-ELM model

3.3 Training and testing process

We split the dataset into two parts: a training dataset of 2400 (50%) samples and a test data set of 2400 samples (50%). We first use the training data to find the best possible parameters for the ELM method. Testing dataset is then used to measure the accuracy of the proposed method. The diagnostic accuracy (Acc) is obtained as

$$Acc = \frac{\text{Total number of hits}}{\text{Total number of samples}} \quad (12)$$

The ELM is calibrated with 6 input nodes and 10 output nodes. The optimal number of hidden nodes is usually very difficult to find. Although it is out of the scope of this paper, in the next paragraph we briefly explain the way we choose the number of nodes in the hidden layer, which is based on a trial and error process.

During the training process, overall hidden weights and bias are randomly initialized. The output weights of the output layer are obtained as explained in Equation (4). Once the training process is done, each ELM model is evaluated with a testing dataset. Figure 2 shows the results obtained for each evaluated model during the testing phase. As we can see, models with 60 and 80 hidden

nodes are not as accurate as the ones that consider 100 hidden nodes. Our results show that the accuracy obtained by those models with 100 hidden nodes varies between 99.968% and 100%. Thus, we decide to build a network that considers 6 input nodes, 100 hidden nodes, and 10 output nodes.

To evaluate the efficiency of the method we propose in this paper, we implement two diagnosis method: a well-known method that combines EMD and SVD [25, 26] and a recently proposed method that combines Fourier coefficient amplitude (FCA) and auto-encoder-ELM [14]. On the one hand, for the EMD-SVD-ELM method, each signal is decomposed using the EMD method [3] and the number of extracted IMFs using this strategy ranges between 9 and 13. In spite of that, the EMD-SVD-ELM model only considers 4 IMF for the input, and 100 hidden nodes (just as in the SWT-SVD-ELM model). On the other hand, for the FCA-AE-ELM method, the Fourier coefficient amplitude of raw vibration signal is computed considering 1024 sub-channels. Since the Fourier transform is symmetric, we only consider the first 512 coefficients, which correspond to the size of the features vector that is passed on to the AE-ELM-classifier.

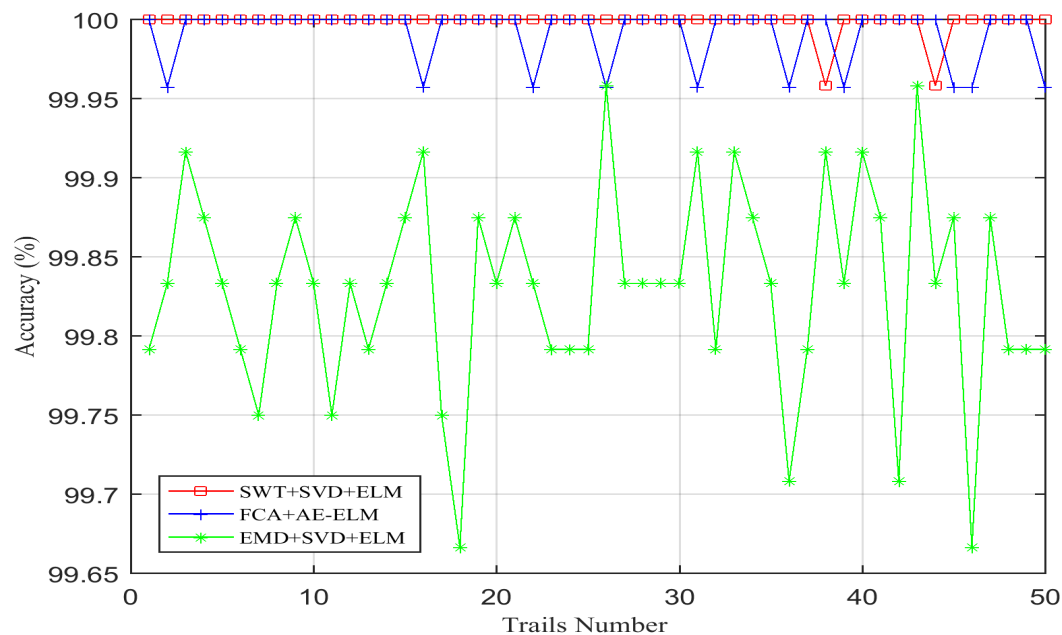


Figure 3. Comparison between the accuracy levels obtained by the SWT-SVD-ELM, EMD-SVD-ELM, and FCA-AE-ELM methods

Results obtained by the all three diagnosis methods implemented in this paper are shown in Figure 3.

As we can see, the accuracy obtained by the EMD-SVD-ELM model ranges between 99.667% and 99.97% with an average accuracy of 99.83%. We point out that this method did not reach the 100% accuracy level for any of the 50 random trials. Unlike the EMD-SVD-ELM, the FCA-AE-ELM does reach the 100% accuracy level in 40 out of 50 random trials. Its minimum accuracy level is 99.957% and its average accuracy is 99.993%. Although better than the EMD-SVD-ELM, the FCA-AE-ELM is still below the SWT-SVD-ELM method we propose in this paper. The SWT-SVD-ELM method reaches the 100% accuracy level in 48 out of 50 random trials. Further, our method obtains an average accuracy of 99.998% and a notable standard deviation of 0.008, which corresponds to a 50% of the standard deviation obtained by the FCA-AE-ELM method (standard deviation of 0.017). This is important as methods with smaller standard deviation values are more stable.

We need to highlight the fact that the average accuracy we obtained using the FCA-AE-ELM model proposed is 99.989%, while the

one reported by the authors in their original paper is only of 98.53% [14]. This is because we modified the topology of the AE-ELM, changing the number of auto-encoders from 3 to 2, where the encoder 1 considers 256 nodes and the encoder 2 considers 126 nodes. The original topology used in [14] considers 650 nodes for the first encoder, 150 nodes for the encoder 2, and 30 nodes for the last encoder.

4. Conclusions

In this paper a method that combines singular values decomposition, stationary wavelet transform, and extreme learning machine for the bearing failure diagnostic problem is presented. The method, we call SWT-SVD-ELM, uses SWT to decompose a vibration signal. Decomposed vibration signal leads to a set of wavelet component signals. SVD is then used to obtain the set of singular values from the corresponding wavelet coefficients matrix. Since singular values remain somehow stable when a change in the rotation velocity occurs, ELM method is able to achieve high levels of accuracy for the failure diagnosis under time-varying speed operation conditions. Based on the experiments performed in this study we found that the best configuration

for the ELM consists of 6 input nodes, 100 hidden nodes and 10 output nodes.

Contributions of this paper are two-fold. First, we have proposed a new diagnosis method that is able to obtain a 100% accuracy level in the 96% of the random trials tested in this paper. This is an important improvement w.r.t. recently proposed methods such as FCA-AE-ELM that only obtains a 100% accuracy in the 80% of the random trials tested here. A second contribution is that we have improved the topology of the AE-ELM proposed in [14] which leads to better accuracy levels: while the topology proposed in [14] obtains an average accuracy of 98.53%, whereas our topology obtains an average accuracy of 99.99%.

As future work we expect to apply our model to run-to-failure lifetime testing. Other topologies for our model can also be investigated in the future.

Acknowledgements

The authors would like to thank the Bearing Data Centre of the Case Western Reserve University for providing the bearing vibration signals datasets and some of the pictures used in this paper.

REFERENCES

1. Abbasion S., Rafsanjani A. Farshidianfar A. & Irani N. (2007). Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine. *Mechanical Systems and Signal Processing*, 21(7), 2933-2945.
2. Coifman R. & Donoho, D.L. (1995). Translation-invariant de-noising. *Wavelets & Statistics: Lecture Notes in Statistics*, 102, pp. 125–150.
3. Huang N.E., Z. Shen & Long S.R. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 454, pp. 903–995.
4. Huang G.B. (2006). Universal Approximation Using Incremental Constructive Feedforward Networks With Random Hidden Nodes. *IEEE Transactions on Neural Networks*, 17(4), pp. 879-892.
5. Huang G.B. & Chen L. (2007). Convex incremental extreme learning machine. *Neurocomputing*, 70(16-18), 3056-3062.
6. Huang G.B., Zhou H., Ding X., & Zhang R. (2012). Extreme Learning Machine for Regression and Multiclass Classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 42(2), pp. 513-529.
7. Daubechies I. (1992). Ten lectures on wavelet, Society for Industrial and Applied Mathematics Philadelphia, PA, USA.
8. Gilles J. (2013). Empirical Wavelet Transform. *IEEE Transactions on Signal Processing*, 61(16), pp. 399-4010.
9. Jinglong C., Zipeng L., Jun Pan, Gaige C., Yanyang Z., Jing Y., Binqiang C. & Zhengjia H. (2016). Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, 70, pp. 1-35.
10. Kedadouche M., Thomas M. & Tahan A. (2016). A comparative study between Empirical Wavelet Transforms and Empirical Mode Decomposition Methods: Application to bearing defect diagnosis. *Mechanical Systems and Signal Processing*, 81, pp. 88-107.
11. Klema V.C. & Laub A.J. (1980). The singular value decomposition: its computation and some applications. *IEEE Transactions on Automatic Control*, 25(2), pp. 164–176.
12. Loparo K.A., Bearing data set, Case Western Reserve University Bearing Data Center, <http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website>.
13. Mallat, S. (1999). A wavelet tour of signal processing, San Diego (CA) Academic Press.
14. Mao W., He J., Li Y. (2017), Bearing fault diagnosis with auto-encoder extreme learning machine: A comparative study. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 231(8), 1560-1578.
15. Mishra C., Samantaray A.K. & Chakraborty G. (2017). Rolling element bearing fault diagnosis under slow speed operation using wavelet de-noising. *Measurement*, 103, pp. 77-86.

16. Nason G. & Silverman, B. (1995). The stationary wavelet transform and some statistical applications. *Wavelets & Statistics: Lecture Notes in Statistics*, 103, pp. 281-300.
17. Pesquet J.C., Krim H. & Carfantan, H. (1996). Time-invariant orthonormal wavelet representations. *IEEE transactions on signal processing*, 44(8), 1964–1970.
18. Purushotham V., Narayanan S., Suryanarayana S. & Prasad A.N. (2005). Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden Markov model based fault recognition. *NDT & E International*, 38(8), pp. 654-664.
19. Qiao H., Zhengjia H., Zhousuo Z. & Yanyang Z. (2007). Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble. *Mechanical Systems and Signal Processing*, 21(2), pp. 688-705.
20. Randall R.B. & Antoni J. (2011). Rolling element bearing diagnostics - A tutorial. *Mechanical Systems and Signal Processing*, 25(2), pp. 485-520.
21. Ruqiang Y., Robert X. & Xuefeng C. (2014). Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal Processing Part A*, 96, pp. 1-15.
22. Serre, D. (2002). *Matrices: theory and applications*, New York: Springer.
23. Tian Y., Jian M., Chen L. & Wang Z. (2015). Rolling bearing fault diagnosis under variable conditions using LMD-SVD and extreme learning machine. *Mechanism and Machine Theory*, 90, pp. 175-186.
24. Wensheng S., Wang F., Zhu H., Zhang Z. & Zhenggang G. (2010). Rolling element bearing faults diagnosis based on optimal Morlet wavelet filter and autocorrelation enhancement. *Mechanical Systems and Signal Processing*, 24(5), 1458-1472.
25. Xiaoyuan Z. & Zhou J. (2013). Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines. *Mechanical Systems and Signal Processing*, 41(1), pp. 127-140.
26. Xiaoyuan Z., Liang Y., Zhou J. & Zang Y. (2015). A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. *Measurement*, 69, pp. 164-179.
27. Yaguo L., Jing L., Zhengjia H. & Ming J. Z. (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 35(1), pp. 108-126.
28. Yaguo L., Zhengjia H. & Yanyang Z. (2011). EEMD method and WNN for fault diagnosis of locomotive roller bearings. *Expert Systems with Applications*, 38(6), 7334-7341.
29. Yang Y., Dejie Y. & Junsheng C. (2006). A roller bearing fault diagnosis method based on EMD energy entropy and ANN. *Journal of Sound and Vibration*, 294(1), pp. 269-277.
30. Yang Y., Dejie Y. & Junsheng C. (2007). A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM. *Measurement*, 40(9), pp. 943-950.
31. Zhang R, Lan Y, Huang G.B. & Xu Z.B. (2012). Universal approximation of extreme learning machine with adaptive growth of hidden nodes. *IEEE Transactions on Neural Networks and Learning Systems*, 23(2), pp. 365-371.