Fault diagnosis of gearboxes using LSSVM and WPT

M. Heidari, H. Homaei[1], H. Golestanian and A.Heidari

Faculty of Engineering, Shahrekord University, Shahrekord, Iran

1Corresponding author: hadi-h@eng.sku.ac.ir

Abstract

This paper concentrates on a new procedure which experimentally recognizes gears and bearings faults of a typical gearbox system using a least square support vector machine (LSSVM). Two wavelet selection criteria Maximum Energy to Shannon Entropy ratio and Maximum Relative Wavelet Energy are used and compared to select an appropriate wavelet for feature extraction. The fault diagnosis method consists of three steps, firstly the six different base wavelets are considered. Out of these six wavelets, the base wavelet is selected based on wavelet selection criterion to extract statistical features from wavelet coefficients of raw vibration signals. Based on wavelet selection criterion, Daubechies wavelet and Meyer are selected as the best base wavelet among the other wavelets considered from the Maximum Relative Energy and Maximum Energy to Shannon Entropy criteria respectively. Finally, the gearbox faults are classified using these statistical features as input to LSSVM technique. The optimal decomposition level of wavelet is selected based on the Maximum Energy to Shannon Entropy ratio criteria. In addition to this, Energy and Shannon Entropy of the wavelet coefficients are used as two new features along with other statistical parameters as input of the classifier. Some kernel functions and multi kernel function as a new method are used with three strategies for multi classification of gearboxes. The results of fault classification demonstrate that the LSSVM identified the fault categories of gearbox more accurately with multi kernel and OAOT strategy.

Keywords*: gearbox, fault diagnosis, wavelet, least support vector machine.*

1. Introduction

To diagnose early fault of rotating machinery, feature extraction of vibration signals is a very important and difficult research task on engineering. The signal of rotating machinery is usually non-stationary, non-linear, and with strong noise interference, meantime, the early signal energy is too low to extract fault features in time domain. Fault diagnosis of Gear box is one of the core research areas in the field of condition monitoring of rotating machines. Diagnosing a gearbox system by examining vibration signals is the most commonly used method for detecting gear and bearing failures. The procedure for a fault diagnosis of a gearbox can be stated in several steps: data acquisition, signal processing, feature selection and diagnostics [1, 2]. To analyze vibration signals, some methods such as time [3, 4], frequency [5], and time–frequency domain [6] have been investigated. Between these, wavelet transform [7, 8, 9, 10] has progressed in the last two decades, and outweighs the other time–frequency ways, although it is lacking in a few aspects as well.

Discrete wavelet transform is primarily considered as an efficient tool for vibration based signal processing for fault detection. Wavelet analysis could provide local features in both time and frequency domains and has the feature of multi-scale, which enables wavelet analysis to distinguish the abrupt components of the vibration signal [11]. The foundations of Support Vector Machines (SVM) have been developed by Vapnik [12, 13] which is applied to both pattern recognition [14, 15, 16, 17, 18] and regression forecasting [19, 20, 21, 22, 23, 24]. The effectiveness of wavelet based features for fault diagnosis of gears using SVM and proximal support vector machines has been revealed by Saravanan et al. [25]. Qu and Zuo [26] utilized a SVM to identify the wear degree of slurry pump. Sun et al. [27] predicted the remaining life of a bearing by establishing a SVR-based model. Hou and Li [28] optimised the parameters of SVR through an evolution strategy and formulated a SVR-based short-term fault prediction strategy. Shen et al. [29] presented a novel intelligent gear fault diagnosis model based on empirical mode decomposition and multi-class transductive support vector machine. Vibration signals extracted from

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]

rotating parts of machineries carries lot many information within them about the condition of the operating machine. Further processing of these raw vibration signatures measured at a convenient location of the machine unravels the condition of the component or assembly under study. Xian and Zeng [30] developed an intelligent fault diagnosis procedure based on wavelet packet transform (WPT) and hybrid SVM. Zamanian and Ohadi [31] presented a method for feature extraction based on exact wavelet analysis to improve the fault diagnosis of gears. In their study, feature extraction was based on maximization of local Gaussian correlation function of wavelet coefficients. They used from a linear support vector machine to classify feature sets extracted with the presented method.

The rest of this paper is outlined as follows. Section 2 briefly describes the fundamental theory of wavelet packet decomposition and two wavelet selection criteria. The proposed new machine health status identification method is presented in Section 3, followed by the experimental verification tests using both bearing and gearbox datasets as stated in Section 4. In Section 5, the effect of different wavelet basis functions on the performance of the proposed scheme is discussed. Conclusions are drawn in Section 6.

2. Theoretical background

2.1. The review of wavelet packet transform

Wavelet packet transform is an extension of discrete wavelet transform. The signals are decomposed into a hierarchical structure of detail and approximations at limited levels as follows:

$$f(t) = \sum_{i=1}^{i=j} D_i(t) + A_j(t) \quad (1)$$

Where $D_i(t)$ denotes the wavelet detail and $A_j(t)$ stands for the wavelet approximation at the jth level [1]. A wavelet packet is a function with three indices of integers *i*, *j* and *k* which are the modulation, scale and translation parameters, respectively,

$$\psi_{j,k}^{i}(t) = 2^{j/2}\psi^{j}(2^{j}t - k), i = 1, 2, 3, \quad (2)$$

The wavelet functions ψ^j are determined as follows:

$$\psi^{2j}(t) = \sqrt{2}\sum_{-\infty}^{+\infty} h(k)\psi^{i}(2t - k) \quad (3)$$

$$\psi^{2j+1}(t) = \sqrt{2}\sum_{-\infty}^{+\infty} g(k)\psi^{i}(2t - k) \quad (4)$$

The original signal *f(t)* is defied after j level of decomposition as follows:

$$f(t) = \sum_{i=1}^{2j} f_j^i(t) \quad (5)$$

While the wavelet packet component signal $f_j^i(t)$ are stated by a linear combination of wavelet packet functions $\psi_{j,k}^{i}(t)$ as follows:

$$f_j^i(t) = \sum_{k=-\infty}^{\infty} c_{j,k}^{i}(t)\,\psi_{j,k}^{i}(t) \quad (6)$$

Where the wavelet packet coefficients $c_{j,k}^{i}(t)$ are calculated by:

$$c_{j,k}^{i} = \int_{-\infty}^{\infty} f(t)\psi_{j,k}^{i}(t)dt \quad (7)$$

Providing that the wavelet packet functions satisfy the orthogonality:

$$\psi_{j,k}^{m}(t)\psi_{j,k}^{n}(t) = 0 \quad if\ m \neq n \quad (8)$$

Two wavelet selection criteria are used and compared to select a suitable wavelet for feature extraction of the problem.

2.2. Maximum relative wavelet energy criterion

Relative wavelet energy gives information about relative energy with associated frequency bands and can detect the degree of similarity between segments of a signal [32, 33]. The energy at each resolution level n, will be the energy content of signal at each resolution is estimated by

$$E(n) = \sum_{i=1}^{m} \left|C_{n,i}\right|^2 \quad (9)$$

Where '*m*' is the number of wavelet coefficients and $C_{n,i}$ is the *i*th wavelet coefficient of nth scale. The total energy can be calculated as follows

$$E_{total} = \sum_{n}\sum_{i} \left|C_{n,i}\right|^2 = \sum_{n} E(n) \quad (10)$$

The distribution of energy probability is defined as follows [30]

$$p_n = \frac{E(n)}{E_{total}} \quad (11)$$

Where $\sum_n p_n = 1$, and the distribution, p_n, is considered as a time scale density. The Total Energy is calculated for each scale and for vibration signals at different rotor speed and for different loading conditions using healthy and faulty gearbox conditions.

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]

2.3. Maximum energy to Shannon entropy ratio criterion

A suitable wavelet is chosen as the base wavelet, which can extract the maximum amount of Energy while minimizing the Shannon entropy of the corresponding wavelet coefficients. The amount of the Energy and Shannon entropy of a signal's wavelet coefficient is shown by Energy to Shannon Entropy ratio [34] and is given as

$$\zeta(n) = \frac{E(n)}{S_{entropy}(n)} \quad (12)$$

In Eq. (12), the entropy of signal wavelet coefficients is given as follows

$$S_{entropy}(n) = -\sum_{i=1}^{m} p_i \,.\log_2 p_i \quad (13)$$

The energy probability distribution of the wavelet coefficients (p_i), is given by

$$p_i = \frac{\left|C_{n,i}\right|^2}{E(n)} \quad (14)$$

With $\sum_{i=1}^{m} p_i = 1$, and $p_i.\log_2 p_i = 0$ if $p_i = 0$.

3. Review of machine learning techniques

3.1. Multi class support vector machine

The SVM is a supervised learning method based on statistical learning theory formulated by Vapnik [12]. The SVM maps the low dimensional data to the high dimensional feature space, and aims to solve a binary problem by searching an optimal hyper plane which can separate two datasets with the largest margin in the high dimensional space. The optimal hyper plane is established through a set of support vectors from the original datasets and these subsets form the boundary between the two classes. The classification function can be described as follows:

$$f(x) = \left(w^T \Phi(x)\right) + b \quad (15)$$

Where the nonlinear mapping function *Φ(x)* maps the input feature vector in to a higher dimensional feature space, *b* is the bias, *w* is the weight vector. *b* and *w* are used to determine the position of the separating hyper-plane. Some problems about multi-class classification have been researched [20, 21]. As seen before, really SVM is a binary classifier. However, rotating machinery may usually suffer more than two faults. To tackle this problem, in this paper three strategies, such as one-against-one (OAO), one-against-all (OAA) and one against others (OAOT) are used [35].

3.2. Least square support vector machine

LSSVM is a reformulation of standard SVM which was proposed by Suykens and Vandewalle [36]. In contrast to SVM, the LSSVM uses a least squares cost function and involves equality constraints instead of inequalities in the problem formulation. Given the training set $\{(x_i, y_i)\}_{i=1}^{n}$ with $x_i \in R^n$ and $y_i \in (-1,1)$. To class the training set, LSSVM has to find the optimal (with maximum margin) separating hyper plane so that LSSVM has good generalization ability. All of the separating hyper planes have the following representation in the feature space: $y(x) = \omega^T \Phi(x) + b$, where ω is the normal vector of the separating hyper plane. Margin maximization is obtained by minimizing the squared norm of ω while also minimizing the fitting error ζ_i of the training set. The resulting optimization problem of LSSVM can be formulated in the following form:

$$\begin{cases} \min j(\omega, \zeta) = \frac{1}{2}\omega^T\omega + \frac{1}{2}\acute{\gamma}\sum_{i=1}^{l}{\zeta_i}^2 \\ subject\ to: y_i[\omega^T\Phi(x_i) + b] = 1 - \zeta_i, i = 1, \dots, l \end{cases} \quad (16)$$

Where $\acute{\gamma}$ is the regularization parameter. The Lagrangian comes in the form:

$$L(\omega, b, \zeta, \alpha) = J(\omega, \zeta) - \sum_{i=1}^{l} \alpha_i\{y_i[\omega^T\Phi(x_i) + b] - 1 + \zeta_i\} \quad (17)$$

Where α_i is the Lagrange multiplier. According to the conditions for optimality yield, the following equations must be satisfied: $\partial L/\partial\omega = 0$; $\partial L/\partial b = 0$; $\partial L/\partial\alpha_i = 0$; and $\partial L/\partial\zeta_i = 0$. Then a linear system for classification and regression can be obtained from the Karush–Kuhn–Tucker conditions [37]. Its solution is found by solving the system of linear equations expressed in matrix form as follows:

$$\begin{bmatrix} 0 & Q^T \\ Q & PP^T + \acute{\gamma}^{-1}I \end{bmatrix}\begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \vec{1} \end{bmatrix} \quad (18)$$

Where

$P = [\Phi(x_1)^T y_1, \dots, \Phi(x_l)^T y_l]$; $\vec{1} = [1, \dots, 1]^T$; and $Q = [y_1, \dots, y_l]^T$; $\alpha = [\alpha_1, \dots, \alpha_l]^T$.

Then the regression function of LSSVM is obtained.

$$f_{LS}(x) = \sum_{i=1}^{l} \alpha_i\, K(x_i, x) + b \quad (19)$$

Where the kernel function can be given by $K(x_i, x) = \Phi^T(x_i).\Phi(x)$ and it meets Mercer's condition. In the process of fault diagnosis, it is very important to choose a reasonable kernel function for support vector machine. Different kernel functions will obtain different decision functions so that determine the operation performance for support vector machine. Generally, two kinds of kernels, i.e. local kernel and global kernel, are utilized to construct the decision functions [38]. A typical local kernel is radial basis function kernel, which is defined as follows

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]

$$K_r(x_i, x) = exp\frac{-(x_i-x)^2}{2\sigma^2} = exp(-\gamma(x_i - x)^2) \quad (20)$$

Where σ is the width of the RBF kernel. A typical global kernel is the polynomial kernel, which is defined as follows

$$K_p(x_i, x) = (x_i^T x + 1)^d \quad (21)$$

Where d denotes the kernel parameter. In order to improve the classification performance and generalization ability for LSSVM, a multi-kernel (K_m) support vector machine (MSVM) is constructed in this study by a controlled parameter β based on the local kernel function K_r and global kernel function K_p.

$$K_m(x_i, x) = \beta K_r(x_i, x) + (1 - \beta)K_p(x_i, x) \quad (22)$$

Where $0 < \beta < 1$ is the controlled parameter. To be an admissible kernel in SVM, kernels must satisfy Mercer's Theorem. Since K_r and K_p all satisfy Mercer's Theorem, therefore a convex combination of them also satisfy Mercer's Theorem. In the MSVM model, there are four parameters: weight parameter β, penalty constant C, kernel parameters σ and d. The weight parameter is used for weight assignment for different kernel function. The penalty constant is used for these samples misclassified by the optimal separating plane and its role is to strike a proper balance between the calculation complexity and the separating error. The kernel function parameters σ and d reflect the characteristics of the training data. All these parameters affect the generalization of MSVM and exert a considerable influence on the performance of MSVM. However, it is not known beforehand which parameters are best for a given problem. In this work, parameters in multi-kernel SVM are randomly selected. The LSSVM was initially proposed to deal with binary classification problems. Multi-classification problems can also be solved by combining a number of binary LSSVMs using any of a number of strategies, such as one-versus-one, one-versus-all and one against others. In this study, OAO, OAA and OAOT methods are used.

4. Experimental validation of the proposed intelligent machine fault diagnosis scheme

Rolling element bearings and gears are the most common and important components used in rotating machinery such as gearboxes. Faults occurring on the surface of these components could cause unexpected machine breakdown. Therefore, it is necessary to develop an effective intelligent gearbox fault diagnosis method. To verify the effectiveness of the proposed method, new gearbox datasets provided by the by Ottawa university in collaboration with the Prognostics and Health Management Society and the test rig experimental setup datasets collected in the Shahrekord University are analyzed.

4.1. Case 1. Ottawa gearbox vibration datasets

Data collected in this section come from Ottawa university gearbox under Prognostics and Health Management Society [39].

Data were sampled synchronously from accelerometers mounted on both the input and output shaft retaining plates of the gearbox. An attached tachometer generates one pulse per revolution providing very accurate zero crossing information. Data were collected at different variable shaft speed under high and low loading. The test runs include seven different combinations of faults and one fault-free reference run. The signals were sampled with sampling frequency 66.666 kHz and the sampling horizon was 4s long.

4.2. Case 2. Shahrekord experimental setup

The experimental setup at Shahrekord University to collect dataset consists of a one-stage gearbox with spur gears, a flywheel and an electrical motor. The test rig has been shown in figure 1. Vibration signals are obtained in the radial direction by mounting the accelerometer on the top of the gearbox. "Easy Viber" data collector and its software, "SpectraPro", are used for data acquisition. The sensitivity and dynamic range of accelerometer probe are 100mv/g and ±50 g. The signals are sampled at 16000 Hz lasting 4 sec. In the present study, four pinion wheels are used. The vibration signal from accelerometer is captured for the following conditions: good gear, gear with tooth breakage, chipped tooth gear and eccentric gear. For bearing vibration signal acquisition five self-aligning ball bearings (1209K) are used. One new bearing is considered as good bearing. In the other three bearings, some defects are created and then various bearings are installed and the raw vibration signals acquired on the bearing housing. So the vibration signals are captured for the following conditions: good bearing, bearing with spall on inner race, bearing with spall on outer race, bearing with spall on ball and bearing with combine defect.

5. Result and discussion

Six base wavelets such as Meyer, symlet16, cofi5, rbio6.8, bior6.8 and db44 are selected for this work. Based on wavelet selection criterion, Daubechies wavelet (db44) and Meyer are selected as the best base wavelet among the other wavelets considered from the Maximum Relative Energy and Maximum Energy to Shannon Entropy criteria respectively. The wavelet packet coefficients of all signals with db44 and Meyer

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]

are calculated at the eighth level of decomposition. After WPT, 2304 statistical features are extracted from the 256 nodes at eight decomposition levels. When applying wavelet transform to a signal, if the Shannon entropy measure of a particular scale is minimum then we can say that a major defect frequency component exists in the scale but, in the present study out of 256 scales considered, the scale having the Maximum Energy to Shannon Entropy of healthy condition is selected, and the statistical features of the wavelet packet coefficient corresponding to the selected level are calculated. Statistical moments like kurtosis, skewness and standard deviation are descriptors of the shape of the amplitude distribution of vibration data, and have some advantages over traditional time and frequency analysis, such as its lower sensitivity to the variations of load and speed. In the present paper, authors' use statistical moments like standard deviation, crest factor, absolute mean amplitude value, variance, kurtosis, skewness and fourth central moment as features to effectively indicate early faults occurring in rolling element bearings and gears. In addition, energy and Shannon entropy of the wavelet coefficients are used as two new features along with other statistical parameters as input of the classifier. These statistical features are fed as input to the soft computing techniques like SVM for fault classification. Two cases of input data and feature sets are considered for classification. In case A, statistical parameters of wavelet packet transform are considered (for each type of the gearbox fault). Case B is related to the condition that statistical features in optimal level, which has been extracted based on the criteria of Maximum Energy to Shannon Entropy ratio, are considered (for each type of gearbox fault). In addition, energy and Shannon entropy factors are used as two new features as features sets in this case. Table 1 shows the results of classification of gearbox with Maximum Energy to Shannon Entropy criterion. In the case B, by Maximum Energy to Shannon Entropy ratio criterion (Table 1), for test set, correctly classified instances for LSSVM is 93.75%. While using 10-fold cross validation average classification accuracy is 92.50% for LSSVM.

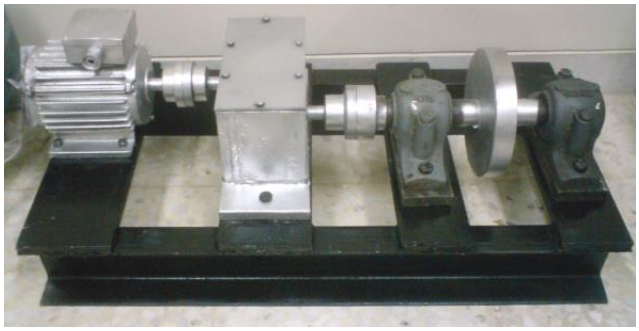


Fig1. Fault simulator set up in Shahrekord University.

Table 1. Classification performance (Maximum Energy to Shannon Entropy criterion)

Parameters		LSSVM (Multi kernel)	
		Test set	10-fold cross validation
Correctly classified	Case A	219 (91.25%)	216 (90%)
	Case B	225 (93.75%)	222 (92.50%)
Incorrectly classified	Case A	21 (8.75%)	24 (10%)
	Case B	15 (6.25%)	18 (7.5%)
Total number of instances		240	240
Training time (s)	Case A (LSSVM)	37.05	
	Case B (LSSVM)	44.72	

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]

Table 2. Classification performance (Maximum Relative Wavelet Energy)

Parameters		LSSVM (Multi kernel)	
		Test set	10-fold cross validation
Correctly classified	Case A	215 (89.58%)	212 (88.33)
	Case B	220 (91.66%)	218 (90.83%)
Incorrectly classified	Case A	25 (10.41%)	28 (11.66%)
	Case B	20 (8.33%)	22 (9.16%)
Total number of instances		240	240
Training time (s)	Case A (LSSVM)	40.94	
	Case B (LSSVM)	43.29	

Table 3. The classified result of experiment data using LSSVM with three methods

Operating condition		Fault classification accuracy based on LSSVM with kernel (%)		
		Polynomial *(d=3)*	RBF*(C=30, γ=2)*	Multi kernel
Outer race fault	OAOT	86.66	86.66	90
	OAA	83.33	86.66	90
	OAO	80	83.33	86.66
Inner race fault	OAOT	90	96.66	96.66
	OAA	86.66	90	93.33
	OAO	86.66	90	90
Roller fault	OAOT	86.66	86.66	90
	OAA	83.33	86.66	90
	OAO	80	83.33	86.66
Combine fault	OAOT	90	90	93.33
	OAA	86.66	86.66	90
	OAO	83.33	86.66	90
Average accuracy (bearing)	OAOT	88.33	89.99	92.49
	OAA	84.99	87.49	90.83
	OAO	82.49	85.83	88.38
Chipped tooth gear	OAOT	90	93.33	96.66
	OAA	90	90	93.33
	OAO	80	83.33	86.66
Eccentric gear	OAOT	90	93.33	93.33
	OAA	86.66	90	93.33
	OAO	83.33	86.66	90
Broken-tooth gear	OAOT	90	90	93.33
	OAA	86.66	86.66	93.33
	OAO	83.3	83.33	90
Good gearbox	OAOT	93.33	93.33	96.66
	OAA	90	90	93.33
	OAO	86.66	83.33	90
Average accuracy (gear)	OAOT	90.83	92.49	94.99
	OAA	88.33	89.16	93.33
	OAO	83.33	84.16	89.16

Table 2 shows accuracy for fault classification with Maximum Relative Wavelet Energy criterion. The correctly classified (case B) instances using test set for LSSVM is 91.66% with two new features. For

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]

10-fold cross validation, average classification accuracies for LSSVM is 90.83%.

From Tables 1 and 2, we found that the Maximum Energy to Shannon Entropy criterion with two new features is better for fault classification of gearbox with respect to Maximum Relative Wavelet Energy criterion. The classification accuracy with LSSVM and Maximum Energy to Shannon Entropy criterion is shown in Table 3. From Table 3, we find that, the classification accuracy with multi kernel by OAOT is better than RBF and polynomial kernels. Figures 3 shows the testing time and training time of LSSVM with three strategies. We can observe that the training time in OAA is bigger than in OAO and OAOT under all kernel functions. It is clear from figure 3, the multi kernel has the least training and testing time with OAOT algorithm. Therefore the OAOT strategy is better than OAO and OAA for the problem. In the case of polynomial kernel, *d* is the important parameter of polynomial kernel, and it is not known beforehand how much value of *d* is the best for classification problem. A 10-fold cross-validation is used to find the best value of *d* and the one with lowest cross validation error is picked. We study the value of *d* from the range *d*={1, 2,… , 8}, the accuracy of three strategies for the multi-class classification is compared in figure 4. From figure 4, we can know that in the case of OAOT algorithm, the accuracy of classification reaches the highest point (89.53%) when *d*=3 and the lowest classification rate as *d*=1. With the grown of parameter *d*, the over-fitting or under-fitting problem is caused and the recognition rate degrades. Generally, the OAOT algorithm is better than OAO algorithm and OAA algorithm under the same value of *d*, and their best classification rate is 82.91% and 86.66%, respectively. Therefore, the optimal result of the polynomial kernel parameter is *d*=3.

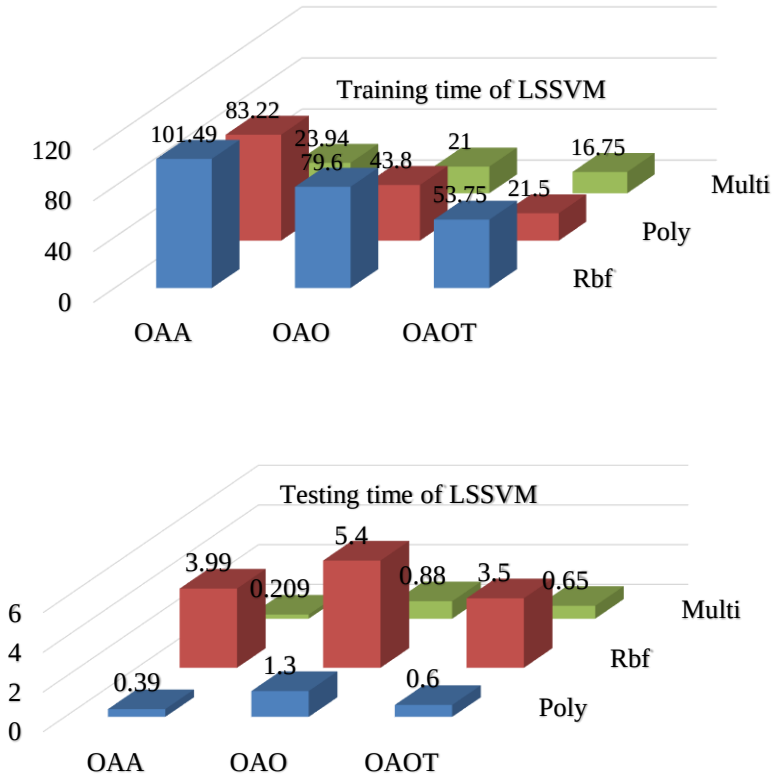


Fig2. Training time and testing time for LSSVM

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]

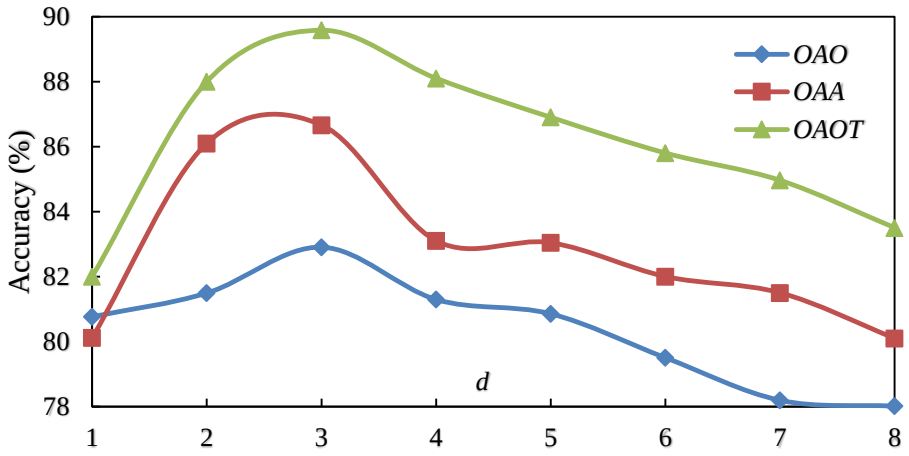


Fig3. Comparison of accuracy of three algorithms based on WPT feature extraction with different d for polynomial kernel

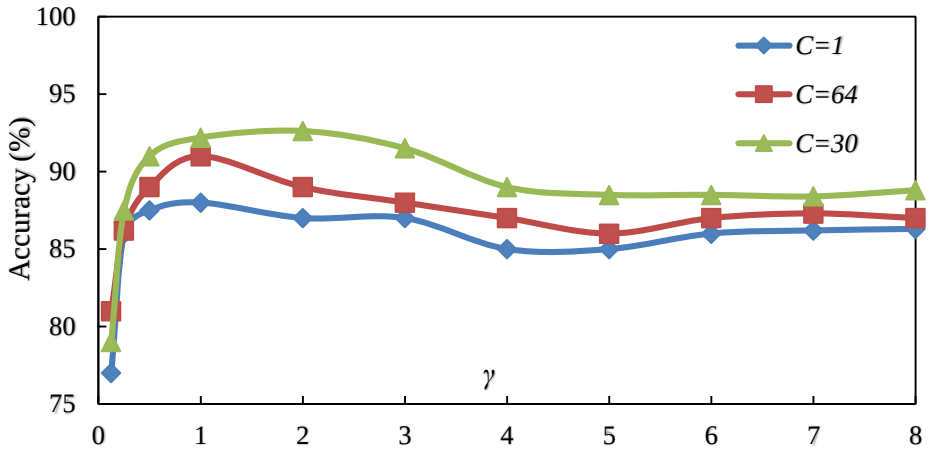


Fig4. Comparison of accuracy using OAOT algorithm based on WPT feature extraction with RBF kernel in different (*C, γ*)

Figure 5 shows that the accuracy of LSSVM using OAOT algorithm with the RBF kernel reaches the highest point (91.24%) with *C=30* and *γ=2*. Similarly, when we apply the RBF kernel to OAO algorithm and OAA algorithm, the best classification ratio is 84.99% and 88.32%, respectively. From Table 3, in the case of multi kernel at LSSVM, we observe that the highest accuracy is 93.74% with OAOT.

6. Conclusions

Fault diagnosis of gearbox is one of the core research areas in the field of condition monitoring of rotating machines. This paper presents, a method for fault classification of a gearbox by LSSVM. First, wavelet packet transform applied over the signal, employing the six mothers wavelet. Two wavelet selection criteria Maximum Energy to Shannon Entropy ratio and Maximum Relative Wavelet Energy are used and compared to select an appropriate wavelet for feature extraction. Results obtained from the two criteria show that the wavelet selected using Maximum Energy to Shannon Entropy ratio criterion gives better classification efficiency. To find very efficient features for classification, Maximum Energy to Shannon Entropy ratio was employed to search for the optimal level decomposition level of wavelet packet and consequently the features were reduced. As a new idea, energy and Shannon entropy have been applied as two new features along with statistical parameters as input of LSSVM. The obtained results indicate that the accuracy of the classifier has been increased between 1 to 4 percentage points by considering these two features but the training time of LSSVM increased with optimal level decomposition and two new features.

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]

Acknowledgements:

The authors are grateful to the Shahrekord University of Iran for supporting the experimental tests of this research. The authors would like to thank Dr. J. Rafiee for his excellent technical help during the experimental aspects and some of the vibration dataset of this work.

Conflict of interests: The authors declare that they do not have any conflict of interests in their submitted paper.

References

[1]. Tran V.T., Yang B.S., An intelligent condition-based maintenance platform for rotating machinery, Expert Systems with Applications, Vol. 39, 2012, p. 2977-2988.

[2]. Melter G., Dien N.P., Fault diagnosis in gears operating under non-stationary rotational speed using polar wavelet amplitude, Mechanical Systems and Signal Processing, Vol. 18, Issue 5, 2004, p. 985–992.

[3]. McFadden P.D., A revised model for the extraction of periodic waveforms by time domain averaging, Mechanical Systems and Signal Processing, Vol. 7, 1993, p. 193–203.

[4]. Combet F., Gelman L., An automated methodology for performing time synchronous averaging of a gearbox signal without speed sensor, Mechanical Systems and Signal Processing, Vol. 21, 2007, p. 2590–2606.

[5]. Minamihara H., Nishimura M., Takakuwa Y., M. Ohta, A method of detection of the correlation function and frequency power spectrum for random noise or vibration with amplitude limitation, Journal of Sound and Vibration, Vol. 141, Issue 3, 1990, p. 425–434.

[6]. Wang W.J., McFadden P.D., Early detection of gear failure by vibration analysis I. Calculation of the time–frequency distribution, Mechanical Systems and Signal Processing, Vol. 3, Issue 7, 1993, p. 193–203.

[7]. Staszewski W.J., Tomlinson G.R., Application of the wavelet transform to fault detection in a spur gear, Mechanical System and Signal Processing, Vol. 8, 1994, p. 289–307.

[8]. Paya B.A., Esat I.I., Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor, Mechanical Systems and Signal Processing, Vol. 11, Issue 5, 1997, p. 751–765.

[9]. Tse P.W., Yang W.X., Tam H.Y., Machine fault diagnosis through an effective exact wavelet analysis, Journal of Sound and Vibration, Vol. 277, 2004, p.1005–1024.

[10]. Wu J.D., Liu C.H., An expert system for fault diagnosis in internal combustion engines using wavelet packet transform and neural network, Expert Systems with Applications, Vol. 36, Issue 3, 2009, p. 4278–4286.

[11]. Cheng J., Yang Y., Yang Y., A rotating machinery fault diagnosis method based on local mean decomposition, Digital Signal Processing, Vol. 22, 2012, p. 356-366.

[12]. Vapnik V., The Nature of Statistical Learning Theory, New York, Springer-Verlag, 1995.

[13]. Cortes C., Vapnik V., Support vector networks, Mach. Learn., Vol. 20, 1995, p. 273–297.

[14]. Bicego M., Figueiredo M.A.T., Soft clustering using weighted one-class support vector machines, Pattern Recognition, Vol. 42, Issue 1, 2009, p. 27–32.

[15]. Cao X.B., Xu Y.W., Chen D., Qiao H., Associated evolution of a support vector machine-based classifier for pedestrian detection, Information Sciences, Vol. 179, Issue 8, 2009, p. 1070–1077.

[16]. Lingras P., Butz C., Rough set based 1-v-1 and 1-v-r approaches to support vector machine multi-classification, Information Sciences, Vol. 177, Issue 18, 2007, p. 3782–3798.

[17]. Zhou S.M., Gan J.Q., Sepulved F., Classifying mental tasks based on features of higher-order statistics from EEG signals in brain–computer interface, Information Sciences, Vol. 178, Issue 6, 2008, p. 1629–1640.

[18]. Zhou S.M., John R.I., Wang X.Y., Garibaldi J.M., Compact fuzzy rules induction and feature extraction using SVM with particle swarms for breast cancer treatments, in: Proceedings of 2008 IEEE Congress on Evolutionary Computation (CEC), 2008, p. 1469–1475.

[19]. Meyer D., Leisch, F., Hornik, K., The support vector machine under test, Neurocomputing, Vol. 55, 2003, pp. 169–186.

[20]. Chuang C.C., Extended support vector interval regress ion networks for interval input–output data, Information Sciences, Vol. 178, Issue 3, 2008, p. 871–891.

[21]. Jayadeva, Khemchandani R., Chandra S., Regularized least squares support vector regression for the simultaneous learning of a function and its derivatives, Information Sciences, Vol. 178, Issue 17, 2008, p. 3402–3414.

[22]. Wong W.T., Shih F.Y., Liu J., Shape-based image retrieval using support vector machines, Fourier descriptors and self-organizing maps,

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]

Information Sciences, Vol. 177, Issue 8, 2007, p. 1878–1891.
[23]. Yuan S.F., Chu F.L., Fault diagnostics based on particle swarm optimization and support vector machines, Mechanical Systems and Signal Processing, Vol. 21, Issue 4, 2007, p. 1787–1798.
[24]. Zhang J., Wang Y., A rough margin based support vector machine, Information Sciences, Vol. 178, Issue 9, 2008, p. 2204–2214.
[25]. Saravanan N., Kumar Siddabattuni V.N.S., Ramachandran K.I., A comparative study on classification of features by SVM and PSVM extracted using Morlet Wavelet for fault diagnosis, Expert Systems with Applications, Vol. 35, 2008, p. 1351–1366.
[26]. Qu J., Zuo M.J., Support vector machine based data processing algorithm for wear degree classification of slurry pump systems, Measurement, Vol. 43, 2010, p. 781–791.
[27]. Sun C., Zhang Z.S., He Z.J., Research on bearing life prediction based on support vector machine and its application, Journal of Physics: Conference Series, Vol. 305, 2011, p. 012028.
[28]. Hou S., Li Y., Short-term fault prediction based on support vector machines with parameter optimization by evolution strategy, Expert Systems with Applications, Vol. 36, 2009, p. 12383–12391.
[29]. Shen Z., Chen X., Zhang X., He Z., A novel intelligent gear fault diagnosis model based on EMD and multi-class TSVM, Measurement, Vol. 45, 2012, p. 30–40.
[30]. Xian G.M., Zeng B.Q., An intelligent fault diagnosis method based on wavelet packer analysis and hybrid support vector machines, Expert Systems with Applications, Vol. 36, 2009, p. 12131–12136.
[31]. Zamanian A.H., Ohadi A., Gear fault diagnosis based on Gaussian correlation of vibrations signals and wavelet coefficients, Applied Soft Computing, Vol. 11, 2011, p. 4807–4819.
[32]. Rosso O.A., Figliola A., Order/disorder in brain electrical activity, Revista Mexicana De Fisica, Vol. 50 , 2004, p. 149–155.
[33]. Rosso O.A., Blanco S., Yordanova J., Kolev V., Figliola A., Schurmann M., Basar E., Wavelet entropy: a new tool for analysis of short duration brain electrical signals, Journal of Neuroscience Methods, Vol. 105, 2001, p. 65–75.
[34]. Yan R., 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, 2007.
[35]. Widodo A., Yang B.S., Support vector machine in machine condition monitoring and fault diagnosis, Mechanical Systems and Signal Processing, Vol. 21, 2007, p. 2560–2574.
[36]. Suykens J.A.K., Vandewalle J., Multiclass least squares support vector machines, in: Proceedings of the International Joint Conference on Neural Networks (IJCNN99), Washington, DC, 2002, p. 900–903.
[37]. Zhao S.L., Zhang Y.C., SVM classifier based fault diagnosis of the satellite attitude control system, International Conference on Intelligent Computation Technology and Automation, 2008, p. 907–911.
[38]. Long B., Xian W., Li M., Wang H., Improved diagnostics for the incipient faults in analog circuits using LSSVM based on PSO algorithm with Mahalanobis distance, Neurocomputing, Vol. 133, 2014, p. 237–248.
[39]. "Data Analysis Competition- 2009", Prognostics and Health management Society. [Online]. Available [Cited: December, 2012]: http://www.phmsociety.org/competition/PHM/09/apparatus.

[Downloaded from i4c2023.iust.ac.ir on 2026-11-21]