Multi-fault Diagnosis of Roller Bearings Using Support Vector Machines with an Improved Decision Strategy

M.M. Manjurul Islam, Sheraz A. Khan, and Jong-Myon Kim[(✉)]

School of Electrical, Electronics, and Computer Engineering,
University of Ulsan, Ulsan, South Korea
{m.m.manjurul, sherazalik, jongmyon.kim}@gmail.com

Abstract. This paper proposes an efficient fault diagnosis methodology based on an improved one-against-all multiclass support vector machine (OAA-MCSVM) for diagnosing faults inherent in rotating machinery. The methodology employs time and frequency domain techniques to extract features of diverse bearing defects. In addition, the proposed method introduces a new reliability measure (SVMReM) for individual SVMs in the multiclass framework. The SVMReM achieves optimum results irrespective of the test sample location by using a new decision strategy for the proposed OAA-MCSVM based method. Finally, each SVM is trained with optimized kernel parameters using a grid search technique to enhance the classification accuracy of the proposed method. Experimental results show that the proposed method is superior to conventional approaches, yielding an average classification accuracy of 97 % for five different rotational speed conditions, eight different fault types and two different crack sizes.

Keywords: Multi-fault diagnosis · Support vector machines · Dempster-Shafer (D-S) theory · Reliability measure · Decision rule

1 Introduction

Reliable fault diagnosis of rolling element bearings (REBs) is a challenging task, especially in industrial machinery. Bearing defects, if not detected in time, can eventually lead to machine failure and cause costly downtime. This has prompted research into vibration analysis [1, 2], motor current signature analysis, and oil debris analysis for fault diagnosis. Vibration analysis has been widely used because it provides essential information about diverse bearing failures [2, 3]. Recently though, the use of acoustic emissions [4] has found prominence in the detection and identification of bearing defects because of its ability to capture low-energy signals that are characteristic of low rotational speeds [4–6]. Specifically, Yoshioka et al. [6] showed that AE could be used to identify bearing defects even before they appear in the range of vibration acceleration, whereas vibration analysis can only to detect defects after they appear on the bearing's surface [3–7].

AE signal based fault diagnosis usually involves fault feature extraction from AE signals, feature selection and fault classification. Fault feature extraction maps the

original signal to statistical parameters that reflect the diverse symptoms of bearing defects. The feature pool in this study includes 10 time-domain statistical parameters, 3 frequency-domain, and 9 complex envelope components. Once we extract features from the AE signal, our problem reduces to achieving better classification accuracy.
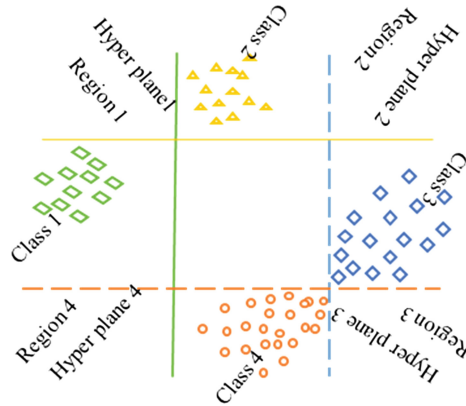
Fig. 1. OAA SVM. Solid lines show true boundary lines and dashed lines linear boundaries

Several classification methods including support vector machines (SVM) [8, 9, 12], artificial neural networks (ANN) [4, 14], fuzzy sets theory, and expert systems have been used in bearing fault diagnosis [3, 5]. SVM is a robust classification tool [13–15] that classifies unknown data into multiple classes [9, 12] by decomposing an L-class into a number of two class problems and constructing a binary classifier for each. In this paper, an improved decision fusion based one-against-all (OAA) multi-class support vector machine is designed for improved fault diagnosis in roller bearings. OAA-MCSVM classifies an unknown feature vector based upon the binary classifier with the largest value for the classifier decision rule. This approach ignores the competence of individual classifiers and relies only on the value of the decision rule, therefore it does not always yield the best results [8–10, 14]. In Fig. 1, which shows the application of the one against all approach to a four class problem, the solid lines separate class 1 and 2 by a definite margin from the remaining classes whereas the dashed lines fail to achieve the same in case of class 3 and 4. Some regions (such as region 3) of the feature space are undecided, where a sample is accepted by more than one class or rejected by all. Individual SVMs are treated equally by the standard OAA-MCSVM method despite differences in their quality, thus compromising the overall classification accuracy. Therefore, it is advantageous to give due weight to each SVM classifier on the basis of classification competence and formulate a new decision strategy to achieve that. The contributions made in this paper are as follows:

- The OAA-MCSVM is modified to improve its classification performance, by introducing a reliability measure, SVMReM, for each SVM and formulating a new decision aggregation rule.

- The effectiveness of the proposed OAA-MCSVM is validated by accurately classify multiple bearing defects under different load conditions and at different rotational speeds.

The remaining parts of the paper is organized as follows. Section 2 explains the experimental setup. Section 3 provides analytical grounds to support our proposition under the framework of Dempster-Shafer (D-S) theory [12] and introduces a reliability measure (SVMReM) for individual SVMs that achieves optimal value irrespective of the test sample location in the feature space. Section 4 describes the proposed method with feature extraction and detailed derivation of reliability measures. Section 5 presents the experimental results, and Sect. 6 concludes this paper.

2 Experimental Setup and Data Acquisition Details

The experimental test rig used for this study is shown in Fig. 2. The AE fault signal is acquired using a general purpose, wideband AE sensor (WS α from Physical Acoustics Corporation) [3–6, 8], placed on top of the bearing housing as shown in Fig. 2. AE sensors are designed to capture high frequency acoustic emissions emanating from periodic impact events involving bearing faults.
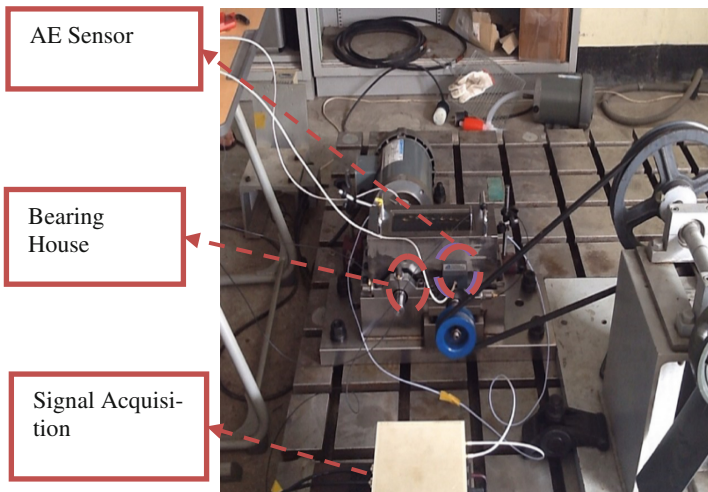
Fig. 2. A self-designed test rig for obtaining AE signals of bearing defects

In this study, one normal and seven different AE fault signals are obtained; (i) (DFN) the normal condition, (ii) (BCO) bearing with outer race crack, (iii) (BCR) bearing with roller crack; (iv) (BCI) bearing with inner race crack, (v) (BCIO) bearing with cracks on inner and outer raceways, (vi) bearing with inner and roller cracks (BCIR), (vii) (BCOR) bearing with roller and outer cracks, and (viii) (BCIOR) bearing with inner, outer, and roller cracks.

3   Reliable One-Against-All (OAA) Multiclass Support Vector Machine (SVM)

The OAA-MCSVM method for an L-class problem creates L binary SVM classifiers to separate each class from the remaining L-1 classes. If we consider a classification problem with P training samples $\{x_1, y_1\}, \ldots, \{x_p, y_p\}$, where $x_j \in R^{Dim}$ is a Dim-dimensional feature vector of the jth training sample and $y_j \in \{1, 2, 3, \ldots, L\}$ is the class to which it belongs. The jth SVM solves the optimization problem given in Eq. (1) [11–14], which provides the jth decision value function.

$$\text{Min } L\left(w, \ \delta_j^k\right) = \frac{1}{2}\|w\|^2 + c\sum_{j=1}^{p}\delta_j^k, subjected, y_j'\left(w_j \cdot \phi\left(x_j\right) + b_j\right) \geq 1 - \delta_j^k, \delta_j^k \geq 0 \tag{1}$$

Where $y_j' = 1$ if $y_j = L, otherwise$ $y_j = -1$. In the classification stage, a test sample x is classified to be in class k such that the decision function Z_k has the highest value as given in Eq. (2).

$$Z_k(x) = (w_k \cdot \varphi(x_k) + b_k), where, k = \mathop{\arg\max}_{k=1,\ldots,L} Z_k(x) \tag{2}$$

3.1   Improving the Standard OAA-MCSVM Using Dempster-Shafer Belief Theory

The proposed approach improves the standard OAA MCSVM using Dempster-Shafer (D-S) evidence theory [12, 16]. Consider a test sample x, a set of L OAA multiclass classifiers SVMk each with a decision function Zk and a set of hypotheses $\Omega = \{L_K\}$ for $1 \leq k \geq L$, where the kth hypothesis asserts that the sample belongs to class k. When each SVMk is applied to x, the result is a part of evidence supporting a certain proposition [8, 10, 15]. We define a basic belief assignment function (BBA) referred to as belief mass mk on hypothesis Ω [10]. It is actually based on the result of the kth classification. When sign (Zk) = 1, it is logical to enhance belief in SVMk proportional to the value of Zk, because "x fits into class k". This piece of evidence alone does not guarantee the truth of hypothesis Lk, only part of it is committed to the belief in hypothesis Lk. The remaining part of the evidence provided by the value of Zk is assigned to the belief in hypothesis Ω as a whole which asserts that "x does not belong to class k". The basic belief assignment (BBA) function mk is defined for an SVM with a positive response.

$$m_k(\{x\}) = \begin{cases} 1 - \exp(-|Z_k(x)| = \lambda_k, if, x = \{L_k\} \\ \exp(-|Z_k(x)| = 1 - \lambda_k, if, x = \Omega \\ 0, else \end{cases} \tag{3}$$

When *sign* (Z_k) = −1, SVM_k classifies x as not belonging to class k. Then, the belief mass function m_k, for such SVMs with negative responses, is defined by (4).

$$m_k(\{x\}) = \begin{cases} 1 - \exp(-|Z_k(x)| = \lambda_k, if, x = \{\overline{L_k}\} \\ \exp(-|Z_k(x)| = 1 - \lambda_k, if, x = \Omega \\ 0, else \end{cases} \tag{4}$$

After finding the BBA values for all the samples and using D-S rule of combination to obtain the combined BBA, the belief function can be computed using Eq. (5) and finally a test sample x is labeled as class k^* with the highest belief.

$$k^* = \begin{matrix} \arg\max \\ k = 1, \ldots, L \end{matrix} Bel(\{L_k\}) = \begin{matrix} \arg\max \\ k = 1, \ldots, L \end{matrix} \lambda_k = \begin{matrix} \arg\min \\ k = 1, \ldots, L \end{matrix} z_k(x)$$
$$= \begin{matrix} \arg\max \\ k = 1, \ldots, L \end{matrix} |z_k(x)|, \tag{5}$$

Similar to Eq. (5), if at least one or more SVM_k generates a positive response, the belief function can be defined by Eq. (6)

$$k^* = \begin{matrix} \arg\max \\ k = 1, \ldots, L \end{matrix} Bel(\{L_k\}) = \begin{matrix} \arg\max \\ z_k(x) \geq 0 \end{matrix} \lambda_k = \begin{matrix} \arg\min \\ k = z_k(x) \geq 0 \end{matrix} z_k(x)$$
$$= \begin{matrix} \arg\max \\ k = 1, \ldots, L \end{matrix} |z_k(x)| \tag{6}$$

However, it is not prudent to place the entire belief in decisions made by the SVMk's. When the independent sources of evidence cannot be fully believed, the BBA should be weakened by a certain factor [16]. Thus an appropriate degree to quantify the amount of belief in every SVMk is proposed and defined in the following section.

4 Methodology

The proposed diagnosis method as shown in Fig. 3, includes feature calculation, reliability calculation of each SVMk, and a new decision strategy for the proposed OAA MCSVMs which are then used for fault classification.

4.1 Feature Calculation

The feature pool is produced based on 10 time domain statistical features, 3 frequency domain features, and 9 envelope spectrum RMS features. Such a rich feature pool minimizes the risk of missing important aspects of the data, and it can achieve greater discrimination among various fault conditions. The time and frequency domain parameters are enumerated in the Table 1, where 'x' is the original time-domain signal with N data samples and "f" is a spectral component of the signal 'x'. We divide 10 s AE signals into different samples of one time period length each. In every revolution, bearing faults give rise to impact signals depending upon their defect frequencies.
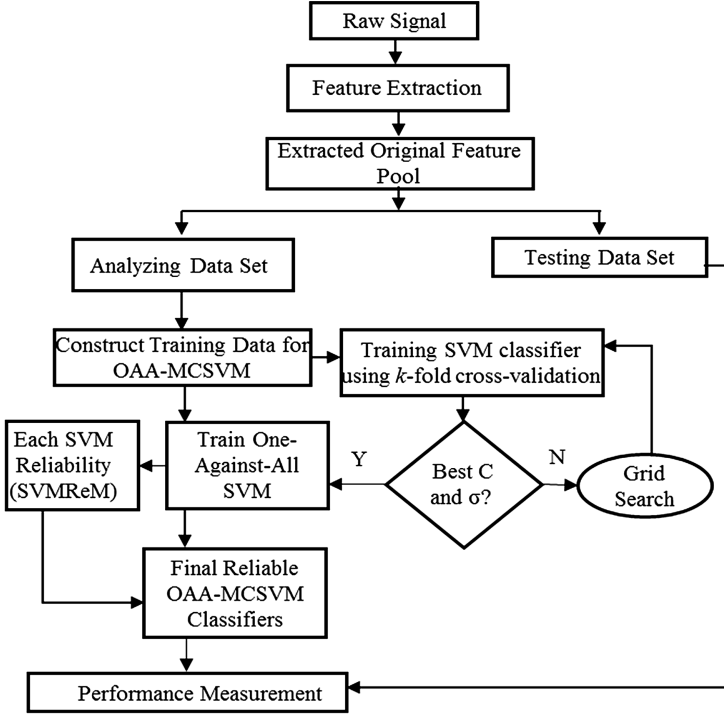
Fig. 3. The proposed fault diagnosis model

Table 1. Multi-domain feature extraction from AE signals

Time-domain feature parameters		
$X_{rms} = \left(\frac{1}{N} \sum\limits_{i=1}^{N} x_i^2 \right)^{\frac{1}{2}}$	$X_{sf} = \dfrac{\max\left(\lvert x_i \rvert\right)}{\left(\frac{1}{N} \sum\limits_{i=1}^{N} x_i^2 \right)^{\frac{1}{2}}}$	$X_{mf} = \dfrac{\max\left(\lvert x_i \rvert\right)}{\left(\frac{1}{N} \sum_{i=1}^{N} \sqrt{\lvert x_i \rvert} \right)^2}$
$X_{kv} = \frac{1}{N} \sum\limits_{i=1}^{N} \left(\frac{x_i - \bar{x}}{\sigma} \right)^4$	$X_{sv} = \frac{1}{N} \sum\limits_{i=1}^{N} \left(\frac{x_i - \bar{x}}{\sigma} \right)^3$	$X_{kf} = \dfrac{\frac{1}{N} \sum\limits_{i=1}^{N} \left(\frac{x_i - \bar{x}}{\sigma} \right)^4}{\left(\frac{1}{N} \sum_{i=1}^{N} x_i^2 \right)^2}$
$X_{ppv} = \max\left(x_i\right) - \min\left(x_i\right)$	$X_{cf} = \dfrac{\max\left(\lvert x_i \rvert\right)}{\left(\frac{1}{N} \sum_{i=1}^{N} x_i^2 \right)^{\frac{1}{2}}}$	$X_{rms} = \left(\frac{1}{N} \sum\limits_{i=1}^{N} \sqrt{\lvert x_i \rvert} \right)^2$
$X_{if} = \dfrac{\max\left(\lvert x_i \rvert\right)}{\frac{1}{N} \sum\limits_{i=1}^{N} \lvert x_i \rvert}$		
Frequency-domain feature parameters		
$X_{fc} = \frac{1}{N} \sum\limits_{i=1}^{N} f_i$	$X_{rmsFr} = \left(\frac{1}{N} \sum\limits_{i=1}^{N} f_i^2 \right)^{\frac{1}{2}}$	$X_{rvf} = \left(\frac{1}{N} \sum\limits_{i=1}^{N} \left(f_i - X_{fc} \right)^2 \right)^{\frac{1}{2}}$

Three types of defects are detected on a bearing depending on their location on the bearing surface. Bearing defect frequency could be measured on the basis of bearing parameters and rotational speed and frequency, for respective defect.

Figure 4 shows the typical signals produced by confined faults in rolling bearing, corresponding envelope signals are calculated using amplitude demodulation [2, 8].
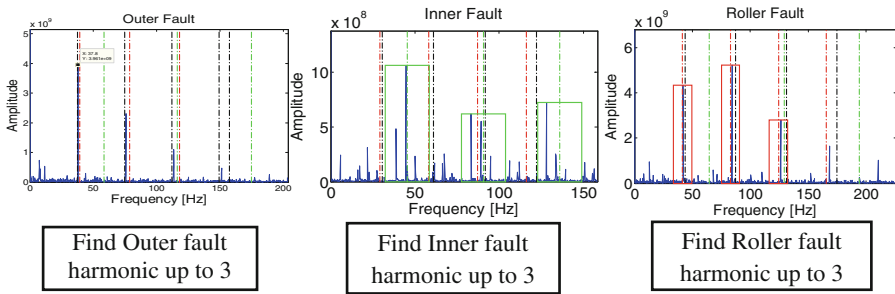
Fig. 4. Complex envelope analysis based on a unique feature extraction method

The defect region for the extraction of features can be determined using Gaussian window method [10, 14], however, it completely ignores the effects of sidebands during defect region calculation [10]. An enveloping spectrum is sensitive to very low impact related events such as sidebands. Thus, we construct a rectangular window around the defect frequency, to overcome problems with Gaussian windows and extract root mean square (RMS) features from the envelope spectrum. A distribution of the extracted features is given in Fig. 5.
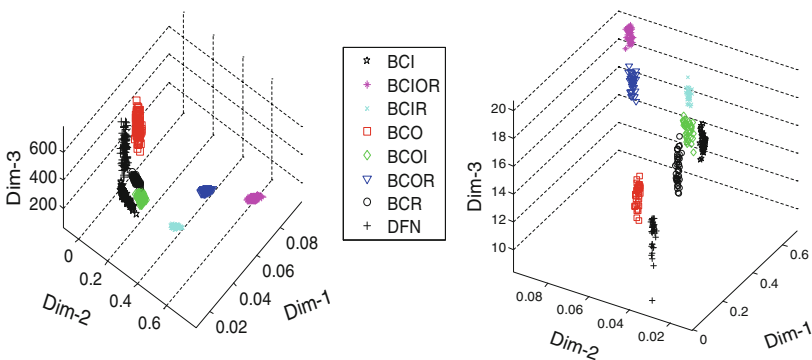
Fig. 5. 3D visualization results of the discriminative fault features: all feature samples of data set #1 (3 MM_300 RPM) for 8 fault conditions (left) and training sample (right)

4.2   Reliability Calculation for Individual SVMs

The effectiveness of a classifier can be calculated using the generalization error R rate, restated as R = E[y = sign (Z(x))], where $y \in \{-1, 1\}$ is the label of true class x and Z is the decision function. A classifier can be measured as more reliable if it provides a smaller value for R [10, 15, 16]. As investigated in [14], when the amount of training sets is relatively small comparing with the dimension of the feature space x, then the small value of Remp will not always ensure a small value regularization error R [11]. In this case, the upper bound of R is defined in [10, 11], and one benefit of SVM is to minimize this objective function value, also minimizes this upper bound of function. In other words, the smaller values of the objective function means smaller regularization errors, and thus it ensures a reliable classifier. So, we restate the objective function for the reliability calculation based on the SVM upper bound optimization equation in (7).

$$ObjFunc = {}^{1}\!/_{2}\left\|\mathrm{w}\right\|^{2} + C\sum_{k=1}^{n}\left(1 - y_k\,z(x_k)\right)_{+} \tag{7}$$

To calculate the ObjFunc, we introduce the unconstrained optimization [9, 10] problem over 'w' in the training stage. Our problem is quadratic; to get a globally optimum solution for 'w', we utilize QP [9]. The objective function would ensure a minimum value of R with the widest margin for all samples in the training stage.

$$\left[\begin{array}{c} \delta \\ {}_{SVM\mathrm{ReM}=\exp\left(-ObjcFunc/\theta\right)=\exp\left(-\dfrac{1/2\|\mathrm{w}\|^2+C\sum_{k=1}^{P}(1-y_k\,z(x_k))_+}{2CP}\right)} \end{array}\right] \tag{8}$$

In order to get the reliability measure for each class, and kernel function in Eq. (8), which yields an optimum value for all samples in the feature space. Furthermore, the margin of the classifier is $\frac{2}{\|w\|}$, a smaller value of $\|w\|$ ensures a larger margin, and thus more accurate generalization. The value $\theta = CP$ is used as a regulation factor to compensate the consequence of C, penalty parameter, P, number of training samples.

4.3   Derivation of Proposed Decision Strategy for Reliable OAA-MCSVM

With the reliability measure $\delta_{SVM\mathrm{ReM}}$ in place, we design a new decision function for our proposed one-against-all multiclass SVM (OAA-MCSVM. First, Zk values of the binary SVMs are calculated for all the given test samples using Eq. (2). Then, a decision fusion is done for all samples. After that, the L class BBA mk is generated based on Eqs. (5) and (6) for all the decisions, where all decisions are considered as information from an independent source. The evidence value that is produced by each classifier SVMk is granted more belief if SVMk is also more reliable, and our calculated competence measures are directly used as a belief factor in Eqs. (9) and (10). The final decision rules as follows:

1. If all the SVM_k produce negative responses,

$$k^* = \begin{array}{c} \arg\min \\ k = 1,\ldots L \end{array} \delta_{SVMReM(k)}\left(1 - \exp^{-|Z_k(x)|}\right) \tag{9}$$

2. If one or more SVM_k produce positive responses,

$$k^* = \begin{array}{c} \arg\max \\ Z_k(x) \geq 0 \end{array} \delta_{SVMReM(k)}\left(1 - \exp^{-|Z_k(x)|}\right) \tag{10}$$

$\delta_{SVM\,Re\,M\,(k)}$ is the reliability measure for SVM_k, and Eqs. (9) and (10) can be merged into one based on Eq. (11).

$$k^* = \begin{array}{c} \arg\max \\ 1,\ldots L \end{array} \left[\delta_{SVMReM(k)}\left\{sign(Z_k(x))\left(1 - \exp^{-|Z_k(x)|}\right)\right\}\right] \tag{11}$$

The value of sign codes the hard decision rule whether "x belongs to class k or not" and the magnitude defines the strength of the decision function value.

4.4 Fault Classification

A SVM discriminates test samples into one of two categories, and subsequently we need multi-class SVMs (MCSVMs) to identify multiple bearing defects. In the OAA multiclass support vector machine approach, each SVM separates one category from the rests, and the final decision is by choosing an SVM that produces the maximum decision value for a given sample. We compare our proposed OAA-MCSVM using a new decision rule with the standard OAA-MCSVM.

5 Experimental Validation

The proposed method is tested on multi-class fault data sets obtained from the experimental setup. The proposed method delivers superior classification performance over its standard counterpart.

5.1 Configuration of the Training and Test Data

The proposed method is evaluated over ten data sets with 90 feature vectors each. Data set for each fault condition, is randomly divided into two subsets for training and testing. The training subset includes 40 randomly-selected feature vectors whereas the remaining 50 feature vectors makeup the testing subset. In the training phase of every SVM, the precision is assessed by several kernel parameters [16] (C = 2^{-5}, 2^{-3}, 2^{-1}..., 2^{15}), and ($\gamma = 2^{-15}$, 2^{-13}..., 2^3) and the best one is chosen using a grid search method

[16] for performance comparison. Consequently, classification accuracy is computed for the standard OAA-MCSVMs and the proposed (reliable) OAA-MCSVMs on the testing datasets; the final classification performance is the average value of the accuracies achieved for each feature vector in the testing dataset.

5.2   Performance Evaluation

To ensure the effectiveness of this proposed OAA-MCSVM method, this experimental analysis compares its classification performance with the standard variant. We use confusion matrix [16] and AUC-ROC based performance analysis for the comparison.

Experiment # 1. This analysis was carried out on ten datasets at 5 rotational speeds (300, 350, 400, 450, 500 RPM), two crack sizes (3 mm and 6 mm) and eight different fault conditions. The support vector machines use the Gaussian radial basis kernel. The results show that our proposed OAA-MCSVM is superior to the standard OAA-MCSVM. The classification accuracy for different data sets is given in Table 2. In addition, the results also validate our unique feature extraction process, which yields classification accuracy of almost 100 % at higher RPMs and larger crack sizes, because

Table 2. Classification performance comparison between standard OAA MCSVMs and the proposed OAA MCSVM for eight fault conditions using radial basis kernel machine (Unit: %)

| Data set	Method	BCI	BCIOR	BCIR	BCO	BCOI	BCOR	BCR	DFN	Model ACC. (%)
3 MM 300 RPM	Standard	97.57	96.57	94.57	98.51	98.91	99.31	99.44	99.17	**98.01**
	Proposed	100	100	96.00	98.93	99.33	99.73	99.87	99.47	**99.17**
3 MM 350 RPM	Standard	97.57	99.44	96.17	98.91	99.57	99.17	97.69	99.57	**98.51**
	Proposed	100	99.43	99.71	99.14	99.86	99.00	99.86	100	**99.63**
3 MM 400 RMP	Standard	99.44	98.77	94.57	99.17	98.91	99.44	99.44	99.57	**98.67**
	Proposed	99.87	99.20	100	99.60	99.33	99.73	99.87	100	**99.70**
3 MM 450 RPM	Standard	99.31	98.64	99.46	99.44	96.57	99.58	99.27	99.59	**98.98**
	Proposed	99.87	99.20	100	99.73	100	99.73	99.73	100	**99.78**
3 MM 500 RPM	Standard	98.57	99.31	99.39	99.64	95.67	100	100	99.57	**99.02**
	Proposed	100	99.73	100	99.87	100	100	100	100	**99.95**
12 MM 300 RPM	Standard	99.57	98.37	97.57	99.57	99.57	99.57	99.57	99.57	**99.17**
	Proposed	100	100	100	100	100	100	100	100	**100**
12 MM 350 RPM	Standard	98.57	99.07	95.67	99.87	96.87	100	100	100	**98.76**
	Proposed	100	100	100	100	100	100	100	100	**100**
12 MM 400 RPM	Standard	99.57	99.57	99.57	99.41	99.67	97.37	99.86	99.52	**99.32**
	Proposed	100	100	100	100	100	100	100	100	**100**
12 MM 450 RPM	Standard	97.97	97.22	99.69	98.59	100	100	100	100	**99.19**
	Proposed	100	100	100	100	100	100	100	100	**100**
12 MM 500 RPM	Standard	98.57	99.97	98.87	99.57	100	98.22	100	100	**99.40**
	Proposed	100	100	100	100	100	100	100	100	**100**

of good separation between fault features. In Fig. 6, we compare the overall performance of Standard OAA MCSVMs and Proposed OAA MCSVMs over ten data sets (in %) using three different kennel functions.

Experiment # 2. We perform the AUC-ROC based analysis of the proposed and standard approaches for all the datasets to verify the robustness of our proposed algorithm. AUC-ROC graph is actually a way of visualization, organization and selection classifiers based on their performance. An area under the receiver operating characteristics (AUC-ROC) analysis has been extended for use behavior analyzing of fault diagnostic systems [14]. In addition, AUC-ROC curves for the proposed method show that it delivers a more uniform performance in the diagnosis of all fault conditions as compared to the standard approach, as shown in Fig. 7.
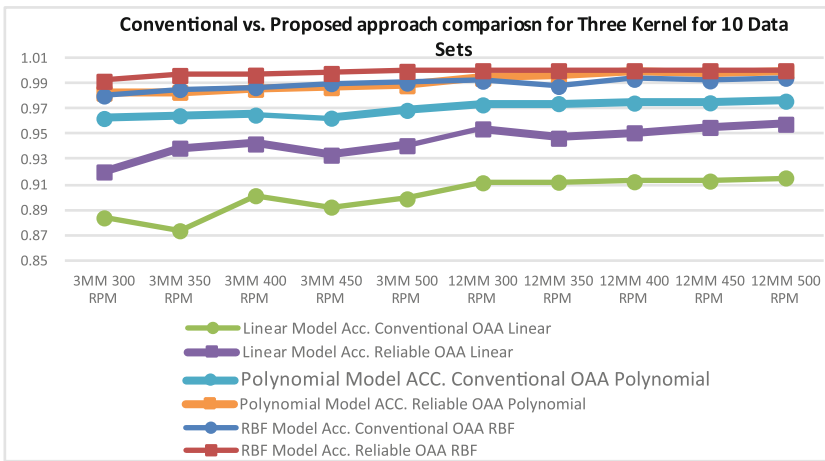
Fig. 6. Relative comparison of the accuracy of proposed and standard methods over three different kennel machines.
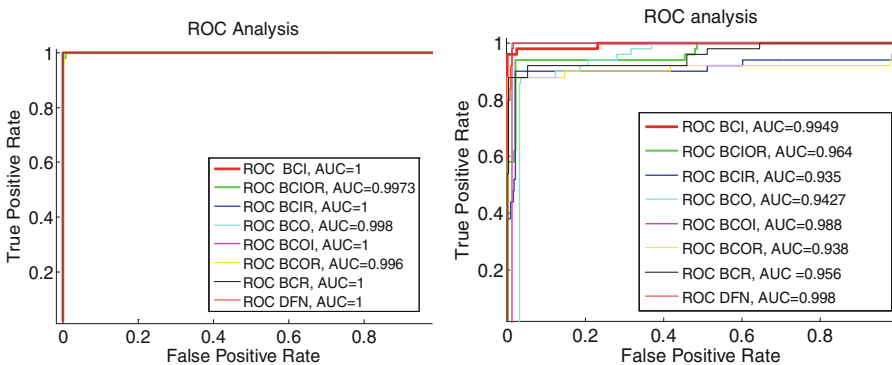
Fig. 7. AUC-ROC characteristic curve of performance comparison between the standard and proposed OAA MCSVM approach.

6 Conclusion

A reliable fault diagnosis methodology for rotating machinery was proposed and evaluated, based upon a modified form of one-against-all multi-class support vector machines, which utilizes individual reliability measures, SVMReM, of the binary SVMs and an improved decision strategy based on the Dempster-Shafer (D-S) theory of evidence. In addition to time and frequency domain analysis techniques, the acquired data from the experiments was preprocessed using envelope analysis methods to extract meaningful features from the fault signals. The experimental analysis demonstrated that the proposed approach yields better classification performance for different fault conditions, at different rotational speeds and different kernel functions for the SVMs. The proposed method yielded an average accuracy exceeding 99 %, 98 %, and 95 % for ten data sets with SVM trained using RBF kernel, polynomial kernel, and linear kernel, respectively.

References

1. Van, M., Kang, H.-J., Shin, K.-S.: Rolling element bearing fault diagnosis based on non-local means de-noising and empirical mode decomposition. Sci. Measur. Technol. IET **8**, 571–578 (2014)
2. Jin, X., Zhao, M., Chow, T.W.S., Pecht, M.: Motor bearing fault diagnosis using trace ratio linear discriminant analysis. IEEE Trans. Ind. Electron. **61**(5), 2441–2451 (2014)
3. Saidi, L., Ben Ali, J., Fnaiech, F., Morello, B.: Bi-spectrum based-EMD applied to the non-stationary vibration signals for bearing faults diagnosis. In: 6th International Conference on Soft Computing and Pattern Recognition, pp. 25–30. IEEE Press (2014)
4. Kang, M., Kim, J., Kim, J.-M.: Reliable fault diagnosis for incipient low-speed bearings using fault feature analysis based on a binary bat algorithm. Inf. Sci. **294**, 423–438 (2015)
5. Ferrando, J.L., Kappatos, V., Balachandran, W., Gan, T.-H.: A novel approach for incipient defect detection in rolling bearings using acoustic emission technique. Appl. Acoust. **89**, 88–100 (2015)
6. Yoshioka, T., Fujiwara, T.: Application of acoustic emission technique to detection of rolling bearing failure. Am. Soc. Mech. Eng. **14**, 55–76 (1984)
7. Harmouche, J., Delpha, C., Diallo, D.: Improved fault diagnosis of ball bearings based on the global spectrum of vibration signals. IEEE Trans. Energy Conserv. **30**, 376–383 (2015)
8. Zhang, X., Chen, W., Wang, B., Chen, X.: Intelligent fault diagnosis of rotating machinery using support vector machine with ant colony algorithm for synchronous feature selection and parameter optimization. Neurocomputing, **167**, 260–279 (2015)
9. Fan, Y., Wang, Z., Zhang, Y., Wang, H.: Fault isolation of non-gaussian processes based on reconstruction. Chemometr. Intel. Lab. Syst. **142**, 9–17 (2014)
10. Soualhi, A., Medjaher, K., Zerhouni, N.: Bearing health monitoring based on Hilbert-Huang transform, support vector machine, and regression. IEEE Trans. Instrum. Measur. Reliab. **64**, 52–62 (2015)

11. Shafer, G.: A Mathematical Theory of Evidence. Princeton University Press, New Jersy (1976)
12. Liu, Y., Zheng, Y.F.: One-against-all multi-class SVM classification using reliability measures. In: IEEE International Joint Conference on Neural Networks, vol. 2, pp. 849–854 (2005)
13. Rauber, T.W., Assis, F., Varejao, F.M.: Heterogeneous feature models and feature selection applied to bearing fault diagnosis. IEEE Trans. Ind. Electron. **62**, 637–646 (2015)
14. Vladimir, N.V.: An overview of statistical learning theory. IEEE Trans. Neural Netw. **10**, 988–999 (1999)
15. Clifton, L., Clifton, D.A., Yang, Z., Watkinson, P., Tarassenko, L., Yin, H.: Probabilistic novelty detection with support vector machines. IEEE Trans. Reliab. **63**, 455–467 (2014)
16. Moreno-Torres, J.G., Saez, J.A., Herrera, F.: Study on the impact of partition-induced dataset shift on -fold cross-validation. IEEE Trans. Neural Netw. Learn. Sys. **23**, 1304–1312 (2012)