Analysis of Statistical Time-Domain Features Effectiveness in Identification of Bearing Faults From Vibration Signal

B. R. Nayana and P. Geethanjali, *Member, IEEE*

***Abstract*— The frequency of rolling element failures in induction motor is high and may lead to losses due to sudden downtime of machine. Researchers are fervent to identify an effective fault diagnosing scheme with less computational burden using optimum number of good discriminating features. We attempted time domain features, namely, waveform length (WL), slope sign changes (SSC), simple sign integral and Wilson amplitude for the first time in addition to established mean absolute value and zero crossing (ZC) for identification of mechanical faults of induction motor. Ten data sets are derived from publicly available vibration database of Case Western Reserve University to identify the capability of features in identification of faults under various conditions. The results are compared with six conventional features for tenfold cross validation using linear discriminant analysis, naive Bayes, and support vector machine classifiers. The results have shown that WL, WAMP, ZC, and SSC outperform other features. Furthermore, area under receiver operator characteristics curve analyses showed an average of 0.9987 with the proposed statistical features and 0.97618 with six conventional features. We also attempted to study the effect of data length and percentage of overlap in classification and found accuracy improves with increase in length but not significant beyond the window length of 3000 with 50% overlap. The proposed statistical features are validated using the brute force method and Laplaician method of feature selection and shown an average accuracy rate of 0.9936 and 0.9894, respectively.**

***Index Terms*— Fault diagnosis, feature extraction, feature selection, pattern classification, statistical features, roller bearings.**

I. Introduction

BEARING faults are more prominently occurring fault [1], hence their diagnosis is crucial in condition monitoring of induction machines. Diagnosing mechanical faults either invasively or non-invasively is more reliable with vibration signals [2]. Over a few decades, several condition monitoring techniques are investigated and there has been substantial advancement with the advantage of expert systems and artificial intelligence algorithms. One of the techniques for condition monitoring of bearing faults is pattern recognition with varying combinations of feature extraction, feature selection, feature reduction and classification. The pattern recognition methods are attempted to diagnose the faults using time domain features, frequency domain features and/or time-frequency domain features from the acquired data. However, the accuracy of fault diagnosis depends on number of features and discriminating capability of the features [3]. Researchers endeavored with time domain features like peak value, crest factor, kurtosis, mean, standard deviation, shape factor and other statistical features [3]–[5] for condition monitoring and found not effective for fault diagnosis. Frequency domain features like power spectrum, power spectral density, periodograms etc. [9]–[10] are attempted. But, these frequency domain features relies on the differences in frequency characteristics of fault conditions [11] and they are not-significant to diagnose. The non-stationary nature of vibration signals have led to time–frequency domain analysis like spectrogram, wavelets transforms (WT) etc. for feature extraction [11]–[16]. These WT analyses suffer a major setback due to adjustable windowed Fourier transform energy leakage during signal processing [16]. Another limitation of this technique is difficulty in selection of appropriate base function to determine the frequency bands of the decomposed signals. Few researchers have extracted many time domain features, frequency domain features and time-frequency domain features to improve the accuracy in identifying the faults [6], [7] with more computational burden. Therefore, a large number of features necessitate the implementation of dimensionality reduction methods prior to classification.

Rauber *et al.* [3] discussed, the addition of unnecessary features deteriorates the performance of classifier. Therefore, authors of this paper attempted to propose simple statistical time domain features to identify the bearing faults with less computational complexity. The classifications are performed with 10 fold cross validation using new time domain features namely, waveform length (WL), slope sign changes (SSC), simple sign integral (SSI), Wilson amplitude (WAMP) along with traditional features mean absolute value (MAV) and zero crossing (ZC). The classification accuracy results are compared with most prominently used six conventional time domain features namely root mean square (RMS), mean, standard deviation (SD), variance (VAR), skewness (SKW) and kurtosis (KURT). The results evince that the proposed time domain features identify the bearing faults with good accuracy compared to conventional features. Further, the accuracy of classification have been investigated with different data length.

Manuscript received February 16, 2017; revised June 22, 2017; accepted July 11, 2017. Date of publication July 26, 2017; date of current version August 9, 2017. The associate editor coordinating the review of this paper and approving it for publication was Prof. Gijs J. M. Krijnen. *(Corresponding author: P. Geethanjali.)*

B. R. Nayana is with the Electrical and Electronics Engineering Department, Sir M Visvesvaraya Institute of Technology, Bangalore, India, and also with the School of Electrical Engineering, VIT University, Vellore 632014, India.

P. Geethanjali is with the School of Electrical Engineering, VIT University, Vellore 632014, India (e-mail: pgeethanjali@vit.ac.in).

Digital Object Identifier 10.1109/JSEN.2017.2727638

TABLE I
DATASET DESCRIPTION

Fault Type	Fault Depth	Dataset Number														
		I		II		III		IV		V		VI	VII	VIII	IX	X
Normal	-	D	D	D	D	D	D	D	D	D	D	D				
Outer race defect	0.007”	D	D					D	D	D	D	D	D	F	DF	
	0.014”			D	D			D	D			D				
	0.021”					D	D	D	D	D	D	D	D	F		DF
Inner race defect	0.007”	D	D					D	D	D	D	D	D	F	DF	
	0.014”			D	D			D	D			D				
	0.021”					D	D	D	D	D	D	D	D	F		DF
Rolling element defect	0.007”	D	D					D	D	D	D	D	D	F	DF	
	0.014”			D	D			D	D			D				
	0.021”					D	D	D	D	D	D	D	D	F		DF
No of loads considered		4	1	4	1	4	1	4	1	4	1	1	1	1	1	1
No of working conditions		16	4	16	4	16	4	40	10	28	7	10	6	6	6	6
No of classes		4	4	4	4	4	4	4	4	4	4	10	6	6	6	6

D-Drive end data, F-Fan end data

In addition, feature selection is investigated using brute force and Laplacian feature selection method. The results from both approaches indicate that the proposed features are ranked higher than conventional features.

The various classification techniques such as k-nearest neighbor (kNN), artificial neural network (ANN) support vector machine (SVM) [15]–[18], linear discriminant analysis (LDA) [6], [7], fuzzy logic [15], [19], [21] and many other combined experts systems are employed for diagnosing bearing faults [5], [7], [13]–[17], [19]–[21]. In this work, authors have used LDA and SVM classifiers and compared the results with newly attempted naive Bayes (NB) classifier. The NB classifier performance with conventional and new statistical features is same. However, proposed features provide maximum accuracy with SVM classifier and the performance of the classifier varies with type as well as number of features. Thus, features selected are validated with area under receiver operator characteristics (AUC-ROC) and accuracy with respect to each dataset.

II. METHODOLOGY

This work examined the publicly available CWRU vibration recordings of 10 seconds from 2 HP, Reliance Electric motor [22]. The drive end (DE) bearing (6205-2RS JEM SKF make) and fan end (FE) bearing (6203-2RS JEM SKF bearing) are seeded with faults using electric discharge machining for depths of 0.007 inch, 0.014 inch and 0.021inch with 0.040 inch of diameter at the inner raceway, rolling element and outer raceway. Accelerometers were placed at 12 O' clock position at DE, FE as well as at baseline of the motor housing to record the vibration signals. The fault depth (FD) in inches will indicate the severity level. The vibration data was recorded with bearing faults for motor load of 0, 1, 2 and 3 HP (motor speeds of 1797 to 1730 RPM). Vibration signals were acquired using a 16 channel DATA recorder, with a sampling rate of 12kHz and 48kHz for drive end bearing faults. Speed and horsepower are measured using torque transducer/encoder were noted. Data recorded at 12kHz is considered for fault identification with respect to DE or FE bearing based on the location of fault for processing in this work.

A. Dataset Description

It is essential to evaluate the effectiveness of the features in various conditions. Therefore, 10 datasets are derived from the CWRU vibration database as shown in Table. 1. The datasets are derived from DE and/FE, bearing condition namely normal (N), inner race defect (IRD), outer race defect (ORD) and rolling element defect (RED) at different fault depths for different load conditions. Dataset I-V, are derived to identify normal, ORD, IRD and RED bearing condition of DE bearing data for all loads to study the effectiveness of the features to identify bearing conditions with respect to fault depths, for individual load condition as well as independent of load. The dataset I-III are derived to perform 4 class classification of faults at fault depth of 0.007” 0.014” and 0.021” respectively. Dataset IV constitutes dataset I-III, to perform classification irrespective of fault severity. The dataset V constitutes dataset I and III. The dataset VI, constitutes 10 different conditions (normal+9(3faults at 3 fault depths)) for individual load. Dataset VII constitutes six classes, to identify fault condition ORD, IRD and RED with respect to fault depth of 0.007” and 0.021” The dataset VIII, is derived similar to dataset VII for FE fault conditions. The dataset IX identifies 6 classes derived from 3 faults of DE and FE bearing for fault depths of 0.007”. The dataset X is similar to IX for fault depth of 0.021”. The number of working conditions for every implementation is product of number of load conditions under consideration and number of fault depths considered as shown in table 1. The number of classes refers to the number of groups derived from working condition for classification. The calculation of working condition and classes is explained with respect to dataset I. The dataset I constitutes normal and 3 faults at fault depth at 0.007” for load conditions of 0HP (L0), 1HP (L1), 2HP (L2) and 3HP (L3). This constitutes a total of 16 working condition. This dataset is grouped for 4 classes namely normal, ORD, IRD and RED for individual load conditions as well as independent of load (4L) condition.

B. Feature Extraction

The statistical time domain features mean absolute value, simple sign integral, waveform length, Wilison amplitude, zero crossing, and slope sign changes are attempted [23], [24] first time for bearing fault diagnosis and compared with conventional features. The features are computed from segment length of L and y_i represents data of the segment. A suitable threshold ϵ is considered for computing WAMP, ZC and SSC.

1) Mean Absolute Value (MAV): The absolute average of data for a segment of length L is defined in equation (1)

$$\mathrm{MAV} = \frac{1}{L}\sum\nolimits_{i=1}^{L} |y_i| \tag{1}$$

2) Simple Sign Integral (SSI): SSI determines the energy of the data segment computed using equation (2)

$$SSI = \sum\nolimits_{i=1}^{L} |y_i|^2 \tag{2}$$

The features MAV and SSI provide amplitude information about the data segment.

3) Waveform Length (WL): WL provides frequency information of data segment calculated from equation (3)

$$WL = WL = \sum\nolimits_{i=1}^{L} |y_i - y_{i-1}| \tag{3}$$

4) Willison Amplitude (WAMP): WAMP is the sum of difference between signal amplitude between two adjacent samples calculated from equation (4). WAMP is calculated if difference exceeds a predefined threshold to reduce noise effects chosen as 0.5.

$$WAMP = \sum\nolimits_{i=1}^{L} f(|y_i - y_{i+1}|) \tag{4}$$

$$f(x) = \begin{Bmatrix} 1 & if\ x \geq \in \\ 0 & otherwise \end{Bmatrix} \tag{5}$$

5) Zero Crossing (ZC): This is another feature provides information about frequency calculated from (6) satisfying equation (7).

$$ZC = (y_i > 0\&\&y_{i+1} < 0)\,||\,(y_i < 0\&\&y_{i+1} > 0) \tag{6}$$

$$|y_i - y_{i+1}| \geq \epsilon \tag{7}$$

6) Slope Sign Change (SSC): SSC is another method to characterize the frequency and computed using equation (8) and (9)

$$SSC = (y_i > y_{i-1}\&\&y_i > y_{i+1})\,||\,(y_i < y_{i-1}\&\&y_i < y_{i+1}) \tag{8}$$

$$|y_i - y_{i-1}| \geq \epsilon \tag{9}$$

7) Root Mean Square (RMS): RMS is amplitude modulated Gaussian random process and is expressed using equation (10)

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{L} |y_i|^2} \tag{10}$$

8) Mean: Mean of a segment indicates the amplitude of the segment and calculated using equation (11)

$$\mathrm{M} = \frac{1}{L}\sum\nolimits_{i=1}^{L} y_i \tag{11}$$

This indicates the first moment of the data considered.

9) Variance (VAR): Variance, indirectly measures the data distribution from mean of the segment. This is the second central moment of a distribution and calculated using equation (12) and (13)

$$VAR = \frac{1}{L}\sum\nolimits_{i=1}^{L} |y_i - \mu|^2 \tag{12}$$

Where

$$\mu = \frac{1}{L}\sum\nolimits_{i=1}^{L} y_i \tag{13}$$

10) Standard Deviation (STD): Standard deviation is the positive square root of variance to measure the variation of data from equation (14) and (15)

$$STD = \sqrt{\frac{1}{L}\sum\nolimits_{i=1}^{L} |y_i - \mu|^2} \tag{14}$$

Where

$$\mu = \frac{1}{L}\sum\nolimits_{i=1}^{L} y_i \tag{15}$$

11) Skewness (SKW): Skewness is the third moment of distribution, to measure the asymmetry of the probability distribution about its mean from equation (16) and (17).

$$\mathrm{SKW} = \frac{\frac{1}{L}\sum_{i=1}^{L} |y_i - \mu|^3}{\left(\sqrt{\frac{1}{L}\sum_{i=1}^{L} |y_i - \mu|^2}\right)^3} \tag{16}$$

Where

$$\mu = \frac{1}{L}\sum\nolimits_{i=1}^{L} y_i \tag{17}$$

12) Kurtosis (KURT): Kurtosis is the scaled form of fourth moment, to find tailness in probability distribution curve and represented using equation (18) and (19)

$$\mathrm{KURT} = \frac{\frac{1}{L}\sum_{i=1}^{L} |y_i - \mu|^4}{\left(\frac{1}{L}\sum_{i=1}^{L} |y_i - \mu|^2\right)^2} \tag{18}$$

Where

$$\mu = \frac{1}{L}\sum\nolimits_{i=1}^{L} y_i \tag{19}$$

C. Feature Selection

Feature selection algorithms are used to improve the accuracy and reduce the computational burden. There are two approaches for feature selection, wrapper approach and filter approach. The wrapper approach identifies features based on classification accuracy and it is classifier dependent. The filter approach uses different benchmarks to critic a feature set and is independent of the classifier. However, the aim of feature selection algorithms is to reduce dimensionality and to improve the performance of classifier. In this paper, feature selection by filter approach is implemented using Laplacian score (LS) and wrapper approach using brute force method. Laplacian method of feature selection is suitable for both supervised and unsupervised forms. LS preserve the locality and based on Laplacian Eigenmaps [25]. The brute force method of feature selection is implemented using 3 classifiers with 10 fold cross validation for possible 63 combinations of the proposed features to identify optimum number of features and the best feature group. The result of best feature group is compared with the best feature group obtained from Laplacian feature selection with same number of features.

D. Classification

In this work an attempt has been made to identify the robustness of the features with the LDA, NB and SVM classifiers. LDA and SVM classifiers are extensively used in literature for diagnosing the bearing faults [26]. SVM is a proven good classifier compared to other conventional classifiers for diagnosing the bearing faults [16], LDA classifier is simple but suffers from the problem of singularity solution [25]–[27]. Therefore, an attempt has been made with NB classifier and compared the performance of bearing fault identification.

The LDA classifier computes projection vector W from class scatter matrix S_w and between class scatter matrix S_b, solving the generalized Eigen value problem for $S_w^{-1} S_b$. The discriminant value is calculated using equation (20).

$$C = Y * W. \tag{20}$$

Where, Y is the n dimensional sample matrix. And C is the transformed matrix. NB is a Bayesian classifier which applies Bayes theorem, and classifies 'm' Classes: $C_1, C_2, C_3 \ldots C_m$.

NB classifier predicts Y belongs to class C_i if and only if following condition is satisfied

$$P\left(C_i/Y\right) > P\left(C_j/Y\right) \quad for\ 1 <= j <= m, j <> i. \tag{21}$$

Where, it follows Naïve assumption of class conditional independence, and X is an 'n' dimensional attribute vector.

SVM is a kernel based classifiers used for linear as well as non-linear classification. SVM classifier maps the data into feature space using kernel functions. In this work, linear SVM is used based on one vs. one approach.

III. Results and Discussions

In this work, the conventional and proposed features are extracted for the datasets described in Table 1. Initially, an attempt has been made to identify the suitable window length and percentage of data overlap for fault identification. From the optimal data segment window, the extracted features are studied using scatter plot. Further, the classification performance of the proposed features and conventional features are studied with and without feature selection techniques. The results are discussed below.

A. Data Segmentation and Feature Extraction

A dataset with maximum number of classes is chosen to quantify the effect of window length with 10-fold cross validations using LDA, NB and SVM classifiers. The conventional and proposed features are extracted from a data window length of 1000, 2000, 3000, 4000 and 5000 samples. The accuracy of classification increases as the window length is increased as shown in Fig. 1(a). The investigation with different window length is performed for every load condition of data set and

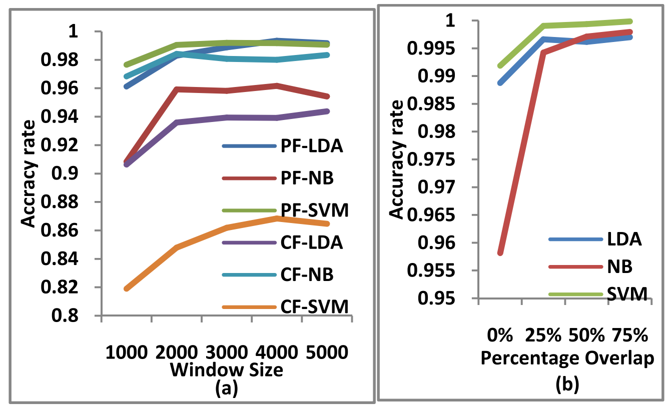

Fig. 1. (a) Illustration of variation in accuracy with respect to window size. (b) Illustration of variation in accuracy with respect to percentage of overlap.

the figure represents average accuracy by using conventional features (CF) and proposed features (PF). It is clear that, improvement in accuracy is insignificant beyond window length of 3000. Therefore, feature evaluations for further studies are performed for a data window size of 3000 samples. Similarly, an investigation on data segment with 0%, 25%, 50% and 75% of data overlap is performed.

Fig. 1(b) represents accuracy rate of different overlap conditions from proposed features. It is observed that accuracy improves with the percentage of overlap. However, improvement is not significant beyond 50% of overlap. Further, 75% of window overlap increases the memory requirement. Thus authors attempted with window size of 3000 with 50% overlap. Therefore, 78 samples are derived for a feature from every working condition. The scatter plot in Fig. 2 demonstrates features of 10 classes considered in dataset VI for no load condition. It is studied from the figure that features WAMP, WL, ZC and SSC have more separability or discriminating capability.

B. Comparison of Proposed Features and Conventional Features

The effectiveness of the features is studied using two feature set (FS), for classification using LDA, NB and SVM classifiers with 10 fold cross validation. The proposed 6 features constitute FS-A and the conventional 6 features constitute FS-B. The percentage of error in classification for datasets I-X is shown in Fig. 3 and Fig. 4. Fig 1(a) represents percentage of error in classification for dataset I with 3 classifiers with both FSs. The results indicate 0% error for any load condition with proposed features and 0-5% error for conventional features.

Similarly the results obtained for other datasets in Fig.1(b)-(f) and Fig 2(a)-(d) can be analyzed and it is clear that proposed features perform better compared to conventional features in most cases.

C. Feature Selection

Brute force method of feature selection with 12 features leads to 4095 combinations for individual classifier. Therefore, authors initially ranked extracted features using Laplacian features selection technique and found the ranking sequence

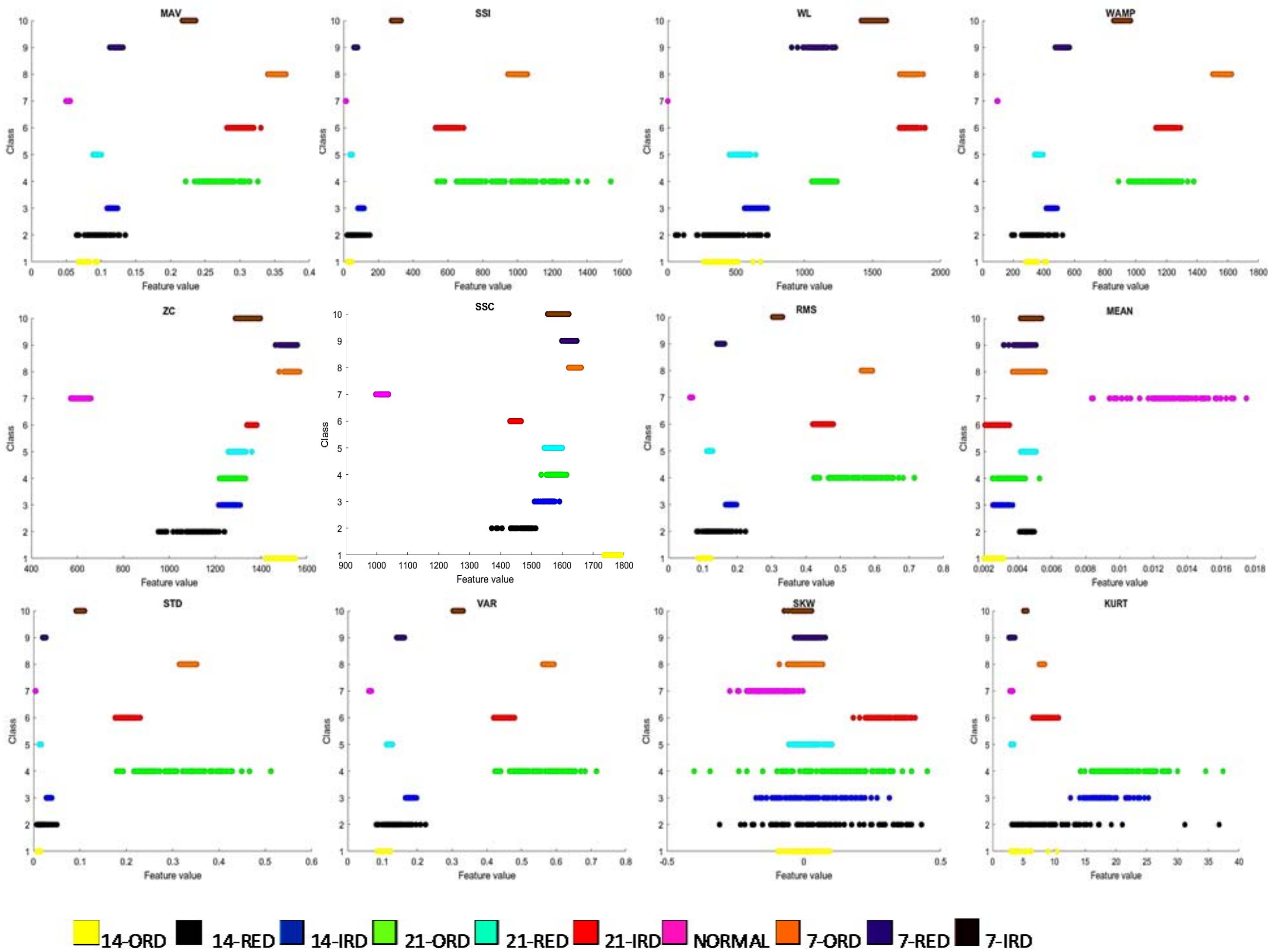

Fig. 2. Scatter plot of 12 features for a dataset of 10 classes.

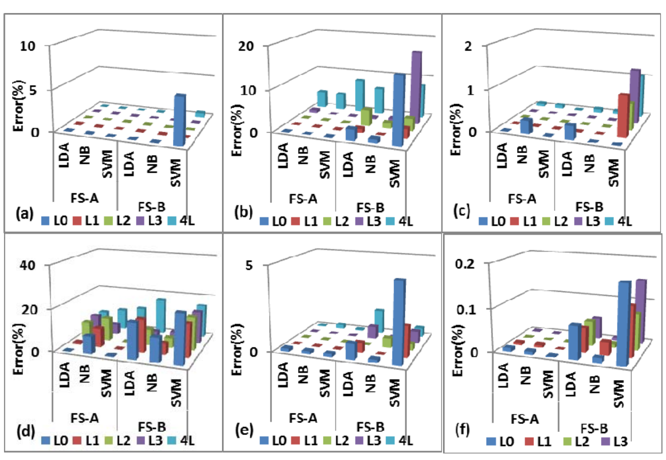

Fig. 3. Accuracy of classification by FS-A and FS-B using LDA, NB and SVM classifiers for datasets I-VI. (a) Dataset-I. (b) Dataset-II. (c) Dataset-III. (d) Dataset-IV. (e) Dataset-V. (f) Dataset-VI.

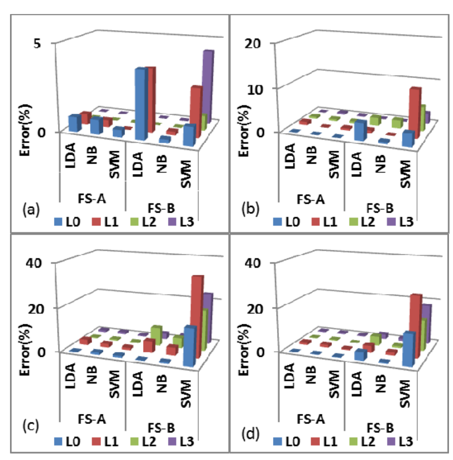

Fig. 4. Accuracy of classification by FS-A and FS-B using LDA, NB and SVM classifiers for datasets VII-X. (a) Dataset-VII. (b) Dataset-VIII. (c) Dataset-IX. (d) Dataset-X.

as {3, 4, 2, 5, 6, 12, 9, 7, 10, 1, 11, 8}. It is clear that the proposed features are ranked high except MAV. Further, optimal number of features and best feature set of proposed features is identified using brute force method of feature selection from 63 combinations with individual classifier. From, brute force method, It has been found that 1-3 varying combination of proposed features are effective in classification for all datasets. The feature set combination resulting maximum accuracy for datasets I-X are {1/2/3/4}, {456}, {356}, {1236}, {1236}, {256}, {256}, {12346}, {456}, {13456}. The results also reveals that, dataset VIII and X exhibited maximum accuracy with 5 features with the negligible difference of 0.056%

TABLE II
CONFUSION MATRIX OF DATASET VI USING LDA CLASSIFIER FOR CF AND PF

CF-SVM		Predicted Labels			
		RED	IRD	ORD	N
Actual Labels	RED	202	31	0	1
	IRD	0	234	0	0
	ORD	70	44	112	8
	N	25	0	0	53

PF-SVM		Predicted Labels			
		RED	IRD	ORD	N
Actual Labels	RED	233	1	0	0
	IRD	0	234	0	0
	ORD	0	0	234	0
	N	0	0	0	78

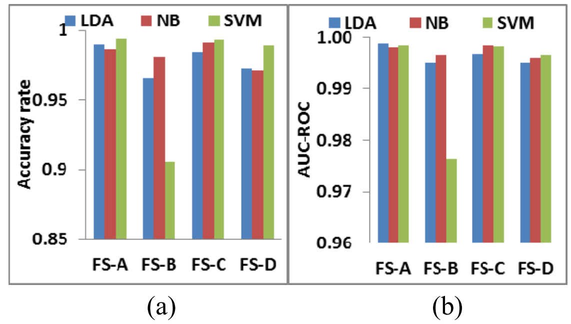

Fig. 5. (a) Average accuracy rate. (b) Average AUC, of all 10 datasets by LDA, NB and SVM for FG-A-D.

while exclusion of MAV feature. Further in dataset I, one of proposed four features results in 100% accuracy. The proposed features ZC and SSC excel in combinations but not individually. Therefore, feature set constituting WL, WAMP, ZC and SSC performs well with good accuracy for all datasets and constitutes FS-C. Subsequently, the first 4 ranked features WL, WAMP, SSI and ZC in Laplacian ranking constitutes FS-D. The accuracy results obtained by FS-D are compared with FS-C for all datasets and found FS-C results in maximum accuracy using SVM classifier. Further, apart from accuracy, effectiveness of the feature or classifier is quantified with AUC-ROC values in the following section.

D. Confusion Matrix and AUC-ROC Analysis

Confusion matrix aids to evaluate the performance and validate the effectiveness of classifier. The confusion matrix of dataset IV of ORD, IRD, RED and normal, the four class classification irrespective of fault severity at no load condition is given in Table II. The confusion matrix of SVM classification is presented for CF and PF and obtained 77.05% and 99.87% of accuracy respectively. Also, it is observed that conventional features are not efficient in diagnosing outer race defect and detects normal and rolling element defect with less error. But, the proposed features have diagnosed the bearing conditions accurately.

Area under curve of ROC, sensitivity versus specificity, graphs for each class describes the effectiveness of features in identification of bearing condition from derived datasets. The area under this curve is unity. if sensitivity versus specificity are unity. These AUC-ROC plots are obtained for each class of the classification and the average AUC-ROC values of all datasets are computed to assist in validating the better performing classifier and feature set. Average AUC and accuracy rates are presented in Fig 5.

From the figure it is inferred that, FS-A, the proposed six features performs better with all three classifiers. The, FS-B of six conventional features has shown poor accuracy rates as well as lower AUC-ROC values. In addition, it is found, the performances vary with classifier chosen. FS-C and D, constitutes a feature set of first four features of brute force ranking and Laplacian ranking respectively. FS-D exhibits variations in accuracy and AUC-ROC values for each classifier and is found to be less than FS-A and C. The classification accuracy and AUC-ROC values is high with FS-A, C and D, but the features selected in brute force method (FS-C) provides good accuracy using SVM classifier.

E. Comparison Study

The comparison of present work with previous works of identical datasets is shown in Table II. The differences exist with respect to sampling rate of data recordings considered that is, 12 kHz or 48kHz and number of input channels taken into consideration. The work of Liu *et al.* [8] used wavelet packet decomposition (WPD) for feature extraction, minimum-redundancy maximum-relevancy (mRMR) for feature selection and differential evolution for classification of dataset I irrespective of load, excluding no load data. Using WPD of 4 levels, 16 features were extracted and 6 were selected from mRMR. The accuracy of 96.1% is achieved using 2 fold cross validation. Similar, work of Amar *et al.* [33] used spectrum imaging and feature enhancement for feature extraction and ANN classification for 2 HP load condition of dataset I and obtained 96.9% accuracy. Wu *et al.* [18] classified dataset III for 0, 1 and 2 HP load using 1-20 features with 99.5-100% accuracy. However, in the present work, 100% accuracy is obtained using 1-3 features for datasets I-III at distinct load conditions. William and Hoffman [20] and Zhang and Li [31] attempted with dataset IV, William and Hoffman [20] employed using characteristics of ZC and classified the faults using feed forward neural networks. In the later work, Zhang and Li [31] attempted using self organizing map along with many classifiers including LDA. The accuracy of 96% and 98% is achieved using 10 time domain and 10 frequency domain features are extracted, and achieved from raw and filtered data using LDA classifier. However, in the present work, an accuracy of 93.14% is obtained from 4 features.

Raj and Murali [21] used fuzzy inference methodology to dataset V without including normal working condition, and obtained a maximum accuracy of 73%. In the present work, 100% accuracy is obtained irrespective of load as well as with respect to each load. Zhang and Zhou [16] studied classification for datasets VII-X using 5 intrinsic mode functions from empirical ensemble mode decompositions. Further, two dimensional feature vector is constructed and classified using SVM classifier. In the present work better or similar results are obtained as shown in Table III. Jin *et al.* [6] classified 10 classes of dataset VI for 3 HP load condition with 9 time domain features and 6 time-frequency features.

TABLE III
COMPARISON OF ACCURACY OF CLASSIFICATION IN LITERATURE AND PRESENT WORK

Reference	Dataset	Load condition	Accuracy (%)	
			Ref Work	Present Work
[8]	I	4L	96.1	100
[33]	I	L-2	96.9	100
[18]	III	L-0	99.56	100
		L-1	100	100
		L-2	99.89	100
[20]	IV	4L	92.5	93.14
[31]	IV	4L	96	
[21]	V	4L	73	100
[16]	VII	L-0	96.81	100
		L-1	97.04	100
		L-2	99.33	100
		L-3	99.7	100
	VIII	L-0	96.9	99.57
		L-1	95.4	98.29
		L-2	98.8	100
		L-3	99.8	100
	IX	L-0	98.17	100
		L-1	98.89	100
		L-2	98.81	99.57
		L-3	98.65	100
	X	L-0	100	100
		L-1	100	100
		L-2	100	100
		L-3	100	100
[6]	VI	L-1	95.95	100
		L-2	95.80	100
[7]		L-3	99.45	100

These features includes skewness, kurtosis and square mean root, shape factor etc. and achieved 99.95% classification accuracy with LDA. In this work, 100% accuracy is achieved using 4 features in LDA, NB and SVM classifiers. Zhao *et al.* [7] did similar study for load conditions of 1HP and 2 HP and obtained 95.95% and 95.80% accuracy respectively with K-mean clustering and LDA classifier. It is clear from the comparative study that proposed statistical features set performs satisfactorily in most of the datasets in diagnosing the bearing faults, with minimum computational burden.

IV. CONCLUSION

In this paper, statistical time domain features MAV, SSI, WL, WAMP, ZC, SSC are attempted for identification of the bearing faults with 10 fold cross validations using LDA, NB and SVM classifiers for 10 datasets derived from CWRU raw data. Investigations are performed to choose the optimal window size and percentage of overlap for feature extraction and found data window of 3000 provides maximum accuracy with 50% overlap. The proposed features are compared with six conventional features RMS, mean, SD, VAR, SKW and KURT and shown the proposed features perform better. Feature selection study using Laplacian feature selection method and brute force method is implemented. Laplacian method ranks features SSI, WL, WAMP, ZC high, and brute force method provides good accuracy with WL, WAMP, ZC and SSC. The average accuracy rate of 0.9894 and 0.9936 is achieved with the feature set combination of Laplacian method and Brute force method. But the average accuracy rate with conventional features is 0.90598. Further, an AUC-ROC curve analysis show an average of 0.9987 with new statistical features and 0.97618 with conventional features. These results evince that proposed features are effective and robust for bearing fault diagnosis of induction machines.

REFERENCES

[1] P. Zhang, Y. Du, T. G. Habetler, and B. Lu, “A survey of condition monitoring and protection methods for medium-voltage induction motors,” *IEEE Trans. Ind. Appl.*, vol. 47, no. 1, pp. 34–46, Jan./Feb. 2011.

[2] F. Immovilli, A. Bellini, R. Rubini, and C. Tassoni, “Diagnosis of bearing faults in induction machines by vibration or current signals: A critical comparison,” *IEEE Trans. Ind. Appl.*, vol. 46, no. 4, pp. 1350–1359, Jul./Aug. 2010.

[3] T. W. Rauber, F. de A. Boldt, and F. M. Varejão, “Heterogeneous feature models and feature selection applied to bearing fault diagnosis,” *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 637–646, Jan. 2015.

[4] J. Chebil, M. Hrairi, and N. Abushikhah, “Signal analysis of vibration measurements for condition monitoring of bearings,” *Australian J. Basic Appl. Sci.*, vol. 5, no. 1, pp. 70–78, 2011.

[5] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, “Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks,” *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3398–3407, Aug. 2013.

[6] X. Jin, M. Zhao, T. W. S. Chow, and M. Pecht, “Motor bearing fault diagnosis using trace ratio linear discriminant analysis,” *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2441–2451, May 2014.

[7] M. Zhao, X. Jin, Z. Zhang, and B. Li, “Fault diagnosis of rolling element bearings via discriminative subspace learning: Visualization and classification,” *Expert Syst. Appl.*, vol. 41, no. 7, pp. 3391–3401, Jun. 2014.

[8] C. Liu, G. Wang, Q. Xie, and Y. Zhang, “Vibration sensor-based bearing fault diagnosis using ellipsoid-ARTMAP and differential evolution algorithms,” *Sensors*, vol. 14, no. 6, pp. 10598–10618, Jun. 2014.

[9] A. Garcia-Perez, R. de J. Romero-Troncoso, E. Cabal-Yepez, and R. A. Osornio-Rios, “The application of high-resolution spectral analysis for identifying multiple combined faults in induction motors,” *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 2002–2010, May 2011.

[10] E. H. E. Bouchikhi, V. Choqueuse, M. Benbouzid, and J. F. Charpentier, “Induction ma chine fault detection enhancement using a stator current high resolution spectrum,” in *Proc. 38th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2012, pp. 3913–3918.

[11] M. Blodt, M. Chabert, J. Regnier, and J. Faucher, “Mechanical load fault detection in induction motors by stator current time-frequency analysis,” *IEEE Trans. Ind. Appl.*, vol. 42, no. 6, pp. 1454–1463, Nov./Dec. 2006.

[12] X. Lou and K. A. Loparo, “Bearing fault diagnosis based on wavelet transform and fuzzy inference,” *Mech. Syst. Signal Process.*, vol. 18, no. 5, pp. 1077–1095, Sep. 2004.

[13] A. Jawadekar, S. Paraskar, S. Jadhav, and G. Dhole, “Artificial neural network-based induction motor fault classifier using continuous wavelet transform,” *Syst. Sci. Control Eng.*, vol. 2, no. 1, pp. 684–690, Dec. 2014.

[14] J. Seshadrinath, B. Singh, and B. K. Panigrahi, “Investigation of vibration signatures for multiple fault diagnosis in variable frequency drives using complex wavelets,” *IEEE Trans. Power Electron.*, vol. 29, no. 2, pp. 936–945, Feb. 2014.

[15] H. L. Schmitt, L. R. B. Silva, P. R. Scalassara, and A. Goedtel, “Classification of simulated motor signals with bearing faults using neural networks, wavelets, and predictability measures,” *IEEE Trans. Ind. Electron.*, vol. 60, no. 2, pp. 567–574, Feb. 2013.

[16] X. Zhang and J. Zhou, “Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines,” *Mech. Syst. Signal Process.*, vol. 41, nos. 1–2, pp. 127–140, Dec. 2013.

[17] B. Samanta, K. Al-Balushi, and S. A. Al-Araimi, “Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection,” *Eng. Appl. Artif. Intell.*, vol. 16, nos. 7–8, pp. 657–665, 2003.

[18] S.-D. Wu, P.-H. Wu, C.-W. Wu, J.-J. Ding, and C.-C. Wang, “Bearing fault diagnosis based on multiscale permutation entropy and support vector machine,” *Entropy*, vol. 14, no. 8, pp. 1343–1356, Jul. 2012.

[19] L. Zhang, G. Xiong, H. Liu, H. Zou, and W. Guo, “Bearing fault diagnosis using multi-scale entropy and adaptive neuro-fuzzy inference,” *Expert Syst. Appl.*, vol. 37, no. 8, pp. 6077–6085, Aug. 2010.

[20] P. E. William and M. W. Hoffman, "Identification of bearing faults using time domain zero-crossings," *Mech. Syst. Signal Process.*, vol. 25, no. 8, pp. 3078–3088, Nov. 2011.

[21] A. S. Raj and N. Murali, "Early classification of bearing faults using morphological operators and fuzzy inference," *IEEE Trans. Ind. Electron.*, vol. 60, no. 2, pp. 567–574, Feb. 2013.

[22] *Case Western Reserve University Bearing Data Center Website*, accessed on Mar. 23, 2015. [Online]. Available: http://csegroups.case.edu/bearingdatacenter/pages/download-data-file/

[23] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Investigating long-term effects of feature extraction methods for continuous EMG pattern classification," *Fluctuation Noise Lett.*, vol. 11, no. 04, p. 1250028, Dec. 2012.

[24] G. Purushothaman and K. K. Ray, "EMG based man–machine interaction—A pattern recognition research platform," *Robot. Auto. Syst.*, vol. 62, no. 6, pp. 864–870, Jun. 2014.

[25] X. He, D. Cai, and P. Niyogi, "Laplacian score for feature selection," in *Proc. Adv. Neural Inf. Process. Syst.*, 2005, pp. 507–514.

[26] K. Fukuaga, *Introduction to Statistical Pattern Recognition*. New York, NY, USA: Academic, 1990.

[27] Y. Elyassami, K. Benjelloun, and M. El Aroussi, "Bearing fault diagnosis and classification based on KDA and alpha-stable fusion," *Contemp. Eng. Sci.*, vol. 9, pp. 453–465, Feb. 2016.

[28] R. K. sharma, V. Sugumaran, H. Kumar, and M. Amarnath, "A comparative study of naïve Bayes classifier and Bayes net classifier for fault diagnosis of roller bearing using sound signal," *Int. J. Decision Support Syst.*, vol. 1, no. 1, pp. 115–129, Sep. 2016.

[29] V. D. Abhijit, V. Sugumaran, and K. I. Ramachandran, "Fault diagnosis of bearings using vibration signals and wavelets," *Indian J. Sci. Technol.*, vol. 9, no. 33, pp. 1–9, Sep. 2016.

[30] T. W. Rauber, F. de A. Boldt, and F. M. Varejao, "Heterogeneous feature models and feature selection applied to bearing fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 637–646, Jan. 2015.

[31] S. Zhang and W. Li, "Bearing condition recognition and degradation assessment under varying running conditions using NPE and SOM," *Math. Problems Eng.*, vol. 4, May 2014, Art. no. 781583.

[32] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," *Mech. Syst. Signal Process.*, vols. 64–65, pp. 100–131, Dec. 2015.

[33] M. Amar, I. Gondal, and C. Wilson, "Vibration spectrum imaging: A novel bearing fault classification approach," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 494–502, Jan. 2015.

[34] E. A. Clancy, E. L. Morin, and R. Merletti, "Sampling, noise-reduction and amplitude estimation issues in surface electromyography," *J. Electromyograph. Kinesiol.*, vol. 12, no. 1, pp. 1–16, Feb. 2002, doi: 10.1016/S1050-6411(01)00033-5.

[35] E. A. Clancy and N. Hogan, "Theoretic and experimental comparison of root-mean-square and mean-absolute-value electromyogram amplitude detectors," in *Proc. 19th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, vol. 3, Oct./Nov. 1997, pp. 1267–1270, doi: 10.1109/IEMBS.1997.756605.

B. R. Nayana received the B.E. degree in electrical and electronics engineering from the Nitte Meenakshi Institute of Technology, Bangalore, India, in 2010, and the M.Tech. degree in signal processing from the S J C Institute of Technology, Bangalore, India, in 2012. She is currently pursuing the Ph.D. degree in electrical engineering with the School of Electrical Engineering, Vellore University, Vellore, India. She was a Research Intern with the Raman Research Institute, Bangalore, from 2011 to 2012. She presented several papers on pulsar de-dispersion.

Since 2013, she has been with the Department of Electrical and Electronics Engineering, Sir M Visvesvaraya Institute of Technology, Bangalore. Her research interests include applications of signal processing in the field of science and engineering.

P. Geethanjali (M'16) received the B.E. degree in electrical and electronics engineering from the Vellore Engineering College, Madras University, Chennai, India, in 2001, the M.Tech. degree in electrical drives and control from the Pondicherry Engineering College, Pondicherry University, Pondicherry, India, in 2004, and the Ph.D. degree from VIT University, Vellore, India, in 2012. Her Ph.D. thesis has been nominated for Best Thesis by the Indian National Academy of Engineering.

Since 2004, she has been with VIT University, Vellore. She received grants from the Department of Science and Technology, Government of India, for research on multichannel myoelectric signal studies for actuation of prosthetic devices and brain controlled wheelchair study for neuromuscular impairments. She also received the Fulbright-Nehru Academic and Professional Excellence Fellowship for 2014–2015. Her research interests include biosignal processing, development of assistive devices, and application of renewable energy in assistive devices.