Machine Learning-based Bearing Fault Classification Using Higher Order Spectral Analysis

Anju Sharma[!,*], G.K. Patra[$] and V.P.S. Naidu[#]

[!]*CSIR-National Aerospace Laboratories, Bengaluru - 560 017, India*
[$]*CSIR-Fourth Paradigm Institute, Bengaluru - 560 037, India*
[#]*Academy of Scientific and Innovative Research (AcSIR), Ghaziabad - 201 002, India*
[*]*E-mail: anjus@nal.res.in*

ABSTRACT

In the defense sector, where mission success often hinges on the reliability of complex mechanical systems, the health of bearings within aircraft, naval vessels, ground vehicles, missile systems, drones, and robotic platforms is paramount. Different signal processing techniques along with Higher Order Spectral Analysis (HOSA) have been used in literature for the fault diagnosis of bearings. Bispectral analysis offers a valuable means of finding higher-order statistical associations within signals, thus proving to detect the nonlinearities among Gaussian and non-Gaussian data. Their resilience to noise and capacity to unveil concealed information render them advantageous across a range of applications. Therefore, this research proposesa novel approach of utilizing the features extracted directly from the Bispectrum for classifying the bearing faults, departing from the common practice in other literature where the Bispectrum is treated as an image for fault classification. In this work vibration signalsare used to detect the bearing faults. The features from the non-redundant region and diagonal slice of the Bispectrum are used to capture the statistical and higher-order spectral characteristics of the vibration signal. A set of sixteen machine learning models, viz., Decision Trees, K-Nearest Neighbors, Naive Bayes, and Support Vector Machine, is employed to classify the bearing faults. The evaluation process involves a robust 10-fold cross-validation technique. The results reveal that the Decision Tree algorithm outperformed all others, achieving a remarkable accuracy rate of 100 %. The naive Bayes algorithm also demonstrated the least performance, with an accuracy score of 99.68 %. The results obtained from these algorithms have been compared with those achieved using Convolutional Neural Network (CNN), revealing that the training time of these algorithms is significantly shorter in comparison to CNN.

Keywords: Bispectrum; Higher-order spectral analysis; Machine learning; Bearings; Condition monitoring

NOMENCLATURE

f_s : Shaft speed (Hz), i.e., $\frac{RPM}{60}$
RPM   : Rotations per minute
D_b : Ball Diameter (mm)
D_c : Cage Diameter (mm)
D_i : Inner Diameter (mm)
D_o : Outer Diameter (mm)
N_B : Number of Balls
θ   : Contact angle (load angle with relation to the radial plane in deg)

1. INTRODUCTION

The condition health monitoring in defense applications, particularly regarding critical components like bearings, is vital. Bearings are the unsung heroes of military machinery, ensuring the smooth operation of aircraft, naval vessels, ground vehicles, missiles, drones, and robotic platforms. Failures in these components can lead to costly downtime, mission failure, and, in worst-case scenarios, compromise national security. Condition health monitoring, specifically using HOSA, is pivotal in safeguarding these assets. As an advanced spectral analysis technique, HOSA offers a unique window into the health of bearings. It excels at capturing subtle high-frequency spectral patterns hidden within vibration signals' patterns that are early indicators of wear, degradation, or impending failure.

Vibration analysis is a prevalent and valuable technique in condition monitoring[1]. Linear analysis may not work for non-linear systems, which show variations before potential failures. The power spectrum is helpful for Gaussian signals but not for non-Gaussian and non-stationary signals in vibration data, which needs phase information. Higher-order signal analysis techniques step in to tackle phase information detection challenges. Among these, the Bispectrum stands out for its ability to process non-Gaussian signals, facilitating the detection and quantification of nonlinearities within time-series data[2] A technique for diagnosing planetary gearbox faults early is suggested[3]. The analysis of modulation signal bispectrum and wavelet packet energy is used in this technique to diagnose defects. Cyclic bispectral slices, frequency shift bispectrum, and bispectral distributions are a few examples of the various bispectral approaches.

Bispectral distribution is recommended as a method of defect diagnostics for the rolling element under circumstances

of changeable load[4]. It was demonstrated that by training Deep Convolutional Neural Networks (CNNs) on images generated from faulty signals' Bispectrum data, highly precise fault detections can be achieved. Various transfer learning approaches were employed, to uncover patterns within the Bispectrum data derived from vibration signals[5]. An approach called the sparse modulation signal bispectrum method has been utilized to determine the bearing's carrier frequency precisely[6]. The current signal has been described utilizing an amplitude modulation feature that combines the higher and lower sidebands simultaneously using a modified bispectrum[7].

Motor faults are detected through motor current signature analysis. This method involves analyzing current signals using both the power spectrum and MSB. Equating the two reveals that MSB exhibits a distinct spectrum structure. MSB outperforms conventional bispectrum in mitigating noise effects on envelope signals[8]. Bispectral characteristics are also of research interest in the field of medicine to identify various disorders. Bispectral-based characteristics have been retrieved to assess the differences between diseased and normal heart sounds[9]. A bispectrum features-based channel selection strategy is suggested in for the brain-computer interface[10]. Similarly, many researchers have employed the Bispectrum in medical fields for disease detection, leveraging its capacity to provide frequency domain information and its inherent ability to suppress Gaussian noise.

In the context of bearing fault classification, traditional approaches have involved treating signals as images using Bispectrum analysis. Notably, the utilization of bispectral features for the classification of bearing faults has not been explored in processing rolling bearing signals[11]. Building upon these observations, our research departs from the established practice of treating signals as images using Bispectrum analysis. Instead, we embrace a distinctive approach, where Bispectrum features were used to classify biomedical signals[9-10]. In this work, features are directly extracted from the Bispectrum for the purpose of classifying bearing faults. This paper is in continuation of our previousresearch paper[12], where we are using the bispectral features for data fusion of two sensors for bearing fault classification. This choice is rooted in the inherent capability of the Bispectrum to provide valuable frequency domain information about bearings, particularly when analysing vibration signals.

Consequently, this paper focuses on delving into and concentrating on the exploration of the frequency band region in the diagonal slice of the bispectrum. This research aims to integrate Bispectrum analysis with machine learning algorithms for health monitoring in rotating machinery. It emphasizes exploring features in the nonredundant region and utilizing fault frequency bands within the diagonal slice of the Bispectrum to enhance fault detection.

1.1 Higher-Order Spectral Analysis

Higher harmonics and their interactions with the signal's phase are discovered, extracted, and studied using higher-order spectra, also called polyspectra. Conventional spectral analysis, which only applies to Gaussian signals, often uses a power spectrum to determine the power distribution among its frequency components. Non-Gaussian signals are analyzed, and signal phase relationships are studied using higher-order spectral analysis. In contrast to power spectrum, it can reveal more details about the non-Gaussian signals.

Bispectrum is the Fourier transform (FT) of third-order statistics. The Bispectrum which isthe 2D Fourier transformation is the signal's third ordercumulant[13]. The third-order cumulants of a zero-mean random process are presented[14-16]. The third-order moment function of $x(k)$ is given by

$$C_{3x}(\tau_1, \tau_2) = E\{x(t)x(t+\tau_1)x(t+\tau_2)\} \tag{1}$$

where, $E\{.\}$ is expectation operator and τ_1 and τ_2 denote the time shift. The cumulants are given by Eqn. (1.2) for a summable third order:

$$\sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} |C_{3x}(\tau_1, \tau_2)| < +\infty \tag{2}$$

Then, Bispectrum of signal $x(t)$ can be defined as:

$$B(w_1, w_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} |C_{3x}(\tau_1, \tau_2)e^{-j(\omega_1\tau_1+\omega_2\tau_2)} \tag{3}$$

The bispectrum using FT for a finite energy signalis given by:

$$B(f_1, f_2) = X(f_1) \cdot X(f_2) \cdot X^*(f_1+f_2) \tag{4}$$

where, B is the Bispectrum of the two frequencies $f1$ & $f2$, and $\omega=2\pi f$ and $X(f)$ is FT of the signal, given by

$$X(f) = \sum_{n=0}^{N-1} x(n)e^{-j\frac{2\pi fn}{N}} \tag{5}$$

$X^*(f)$ being the complex conjugate of $X(f)$. The Bispectrum analysis explores the relationship between sinusoids at fundamental frequencies (f_1, f_2) and their modulation at (f_1+f_2), forming a triplet of frequency components. It calculates the correlation between these components.

2. ESTIMATION OF BISPECTRUM

The bispectrum estimation can use direct or indirect methods. The direct approach involves computing the discrete Fourier transform of the signal's third-order autocorrelation, which is simpler, and this paper adopts this method.

2.1 Direct Method[16]

- Split the signal into K segments of M samples. The samples can be $x(1)$, $x(2)$……$x(k)$ ensuring $M=2^l, l \in Z^+$ for FFT convenience.
- Apply DFT to each of the K segments i.e
  $$X^i(\lambda) = FFT(x^i(k)) \tag{6}$$
- Determine each segment's bispectrum from the DFT coefficients using
  $$b^i(\lambda_1, \lambda_2) = X^i(\lambda_1)X^i(\lambda_2)X^i(\lambda_1+\lambda_2) \tag{7}$$

- Calculate bispectrum's average estimate over all the segments using
  $$B_3^x(\omega_1, \omega_2) = \frac{1}{K}\sum_{i=1}^{K} b^i(\omega_1, \omega_2) \tag{8}$$

where, $\omega = (\frac{2\pi f}{N_0})\lambda$ and N_0 is total samples in a segment.

3. METHODOLOGY

This work uses the data obtained from the oscillatory bearing test rig shown in Fig. 1 which has been installed at CSIR-NAL to investigate bearing defects. In this study ball

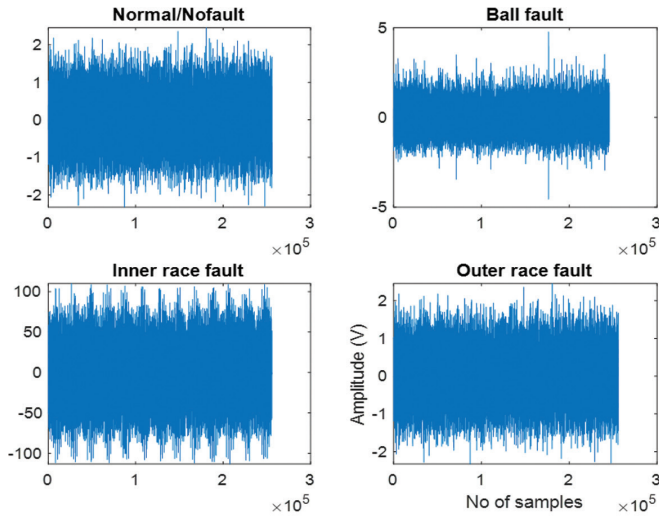**Figure 1. Oscillatory bearing test rig at CSIR-National Aerospace Laboratories.**

Figure 2. Raw data of normal/no fault, ball fault, Inner fault, and outer fault defects.

bearing with a single row of ballsis used for testing. Three different types of defects: ball defect, inner race defect, and outer race defect have been manually seeded. Various loads, ranging from 0 to 40 kgs in 10 kgs increments, were applied to evaluate its performance. Figure 1 illustrates the testing rig utilized at CSIR-NAL for bearing assessment. A vibration sensor (Type 352C33) is placed in the vertically on the housing where bearing is placed. The data was collected from the vibration sensor at a sampling rate of 51200 samples/sec with a load of 40 kg and a rotational speed of 1500 rpm. During the bearing test, four different signals were acquired: Normal data (fault-free), Ball defect data, Inner race fault data, and outer race fault data. Figure 2 displays the raw signals of all the faults. These raw vibration signals are segmented into 256 samples, from which the Bispectrum is obtained. Figure 3 shows the proposed methodology applied in this research[12].

4. NON REDUNDANT REGION AND DIAGONAL SLICE OF BISPECTRUM

Due to the conjugate symmetry of the FT, the bispectrum exhibits symmetrical properties. So, the bispectrum is
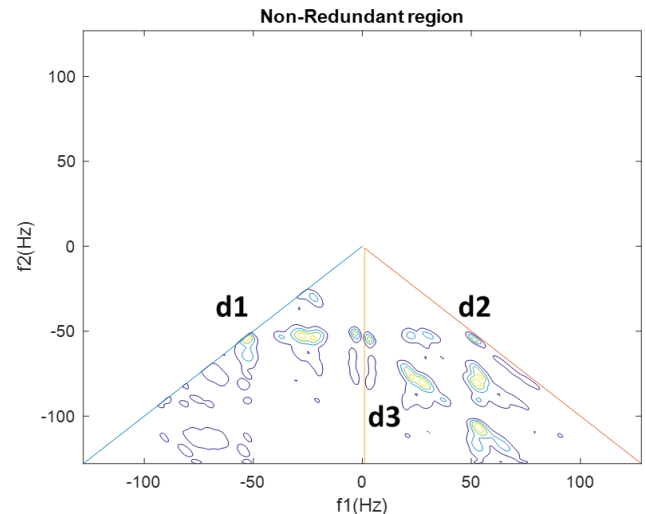

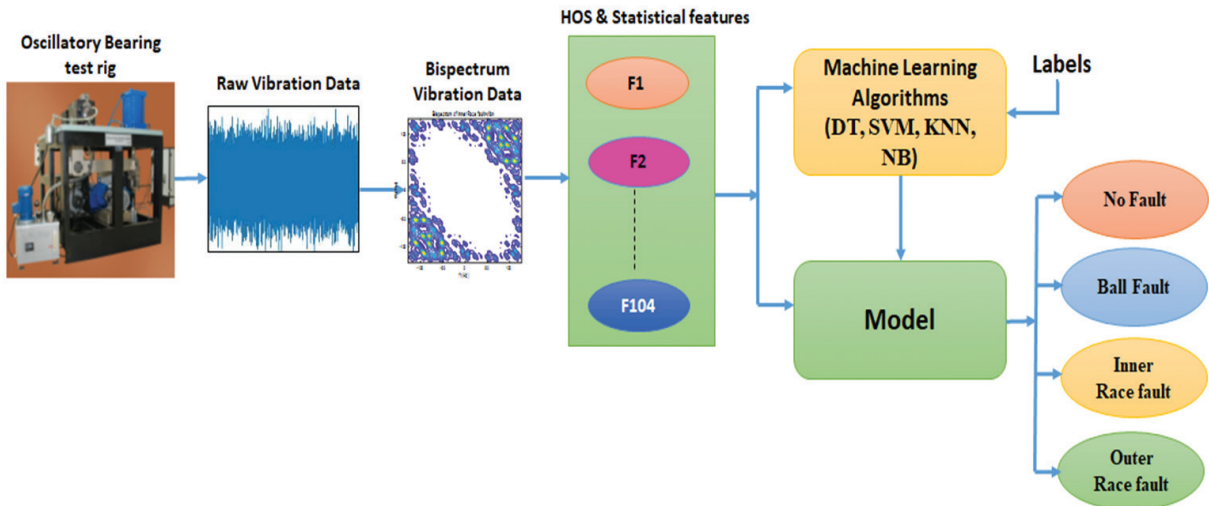**Figure 4. Non Redundant region of outer race fault.**

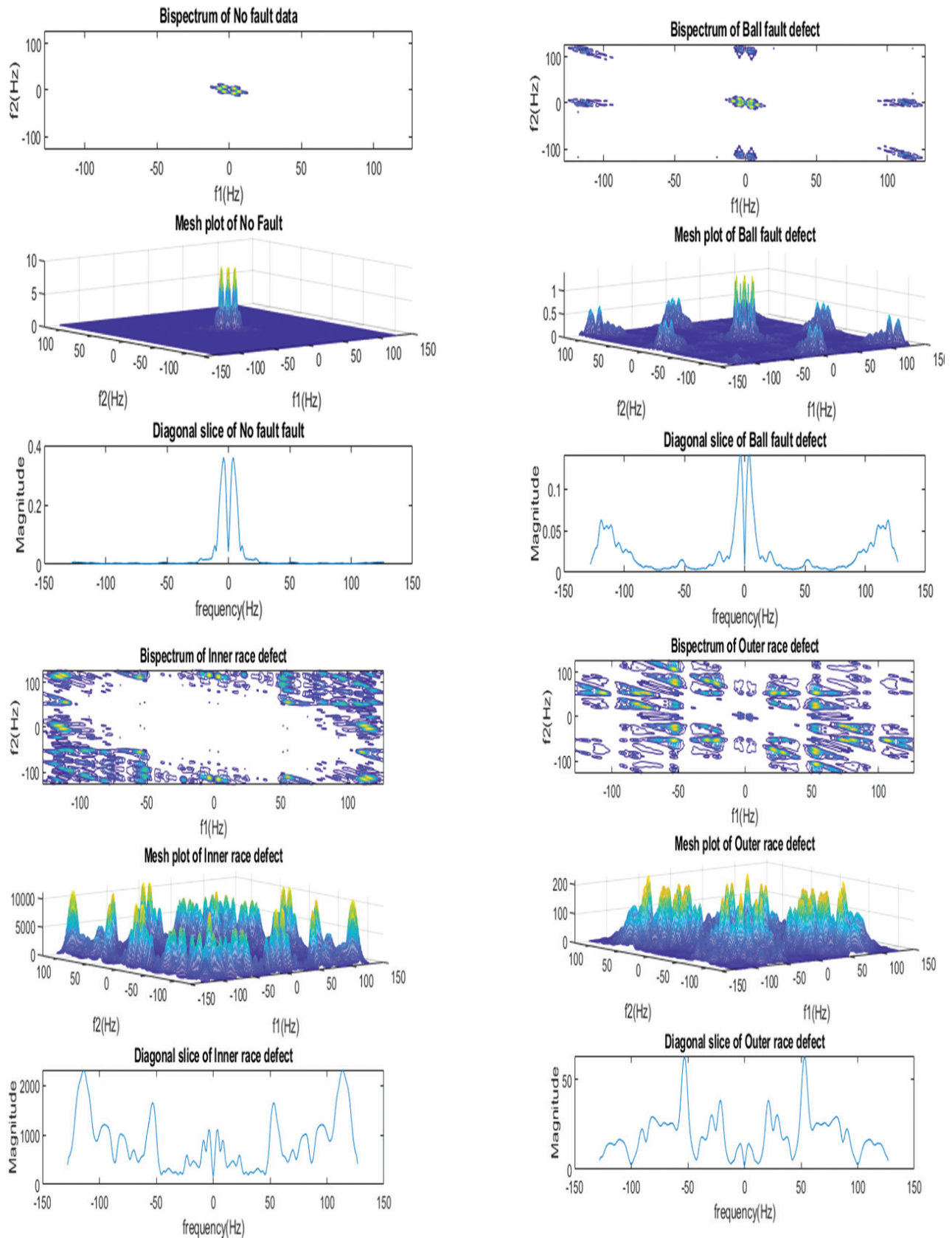**Figure 3. Block diagram of the proposed methodology.**

Figure 5. Bispectrum, mesh plots and diagonal slices of normal data, ball fault, inner race fault and outer race fault.

considered in the triangular region known as the non-redundant region (Principal domain), Ω[9-10,14-15]. Figure 4 depicts the non-redundant region for the outer race fault for 256 samples. The

diagonal elements in the non-redundant region are d1 and d2, and d3 is the line that divides it into two halves.

The one-dimensional FT of the third-order cumulants is called the diagonal slice of the signal[15], which can be defined as

$$B_s(f) = \sum_{\tau=-\infty}^{\infty} C_{3x}(\tau)e^{-j2\pi ft} \tag{9}$$

The bispectrum is projected onto the plane $f_1=f_2$ to create the bispectrum diagonal slice. This approach is computationally more efficient than other HOSA because the diagonal slice offers the advantage of reduced computational demands highlighting peak frequencies. Figure 5 displays the Bispectrum, Mesh plots, and diagonal slices for Normal data, Ball fault, Inner race fault, and Outer race fault. These plots reveal variations in the bispectrum when different faults are present. Normal data exhibits spectral peaks at the center of the contour plots. In contrast, Ball Fault data shows central peaks surrounded by side peaks. Notably, inner-race faults exhibit a more pronounced and stronger central peak compared to outer-race faults with larger but less intense peaks. These variations underscore the usefulness of bispectrum characteristics.

The diagonal slice of the bispectrum, represented in FFT terms by Eqn. (9), enhances the fundamental frequency F_0[15-16]. Due to the diagonal slice's symmetrical properties, frequencies are mirrored on the right-hand side of 0 Hz, requiring consideration of only one frequency side.

$$B(f) = [X(f)X(f)X(2f)] \tag{10}$$

5. FEATURE EXTRACTION

In the principal domain/ non-redundant portion of the bispectrum, only the phase and frequency-linked bifrequency peaks can be observed. On the other hand, the FT of the diagonal slice exhibits peaks at each of the frequencies implicated in the phase coupling. The fault bands about the fundamental fault frequencies are calculated, and the following features are extracted from fault bands: sum of all peaks in a frequency band, average and maximum of all peaks in the principal diagonal sum of all frequencies of principal diagonal d1, d2, and d3 of the non-redundant region. Annexure I shows the features F1-F35 extracted from the non-redundant region, and Table 3 shows the features extracted from the fault frequency bands. The final step is integrating all 104 features and using them for fault classification

.

5.1 Fundamental Fault Frequencies

The fundamental fault frequencies are calculated from the diagonal slice of the bispectrum for all the bearing faults, i.e., Normal data, Outer race fault, inner race fault, and ball fault. The conventional formulae[17] presented in Eqn. (11-14) compute the bearing fundamental fault frequencies assuming the inner race is rotating and the outer race is stationary.

f_{BPOF}	Ball pass outer case frequency	$\dfrac{N_B}{2}f_s\left(1-\dfrac{Db}{Dc}\cos\theta\right)$	(11)
f_{BCF}	Ball cage frequency	$\dfrac{f_s}{2}\left(1-\dfrac{Db}{Dc}\cos\theta\right)$	(12)
f_{BPIF}	Ball pass inner case frequency	$\dfrac{N_B}{2}f_s\left(1+\dfrac{Db}{Dc}\cos\theta\right)$	(13)
f_{BSF}	Ball spin frequency	$\dfrac{D_c}{2D_b}f_s\left(1-\dfrac{Db^2}{Dc^2}\cos^2\theta\right)$	(14)

Fault frequencies like *Fx* and its harmonics *2Fx*, *3Fx*…., etc. often appear as peaks in the frequency spectrum. However, interference from other vibrations can make them hard to detect. Additionally, variations in speed and load levels may cause deviations from their expected values, posing extraction challenges. Bearings usually operate within a 0-6 kHz frequency range. Figure 6 shows the bearing geometry for fault frequency calculations, emphasizing the need for Cage diameter accuracy. Frequency calculations require the cage diameter D_c, which can be estimated from the Eqn. (15)

$$D_c \approx \frac{D_b}{2} + D_i + \frac{D_b}{2} \approx 63mm \tag{15}$$

The dimensions of each component of the bearing in mm are given as follows:

$$\theta = 0$$
$$N_b = 10$$
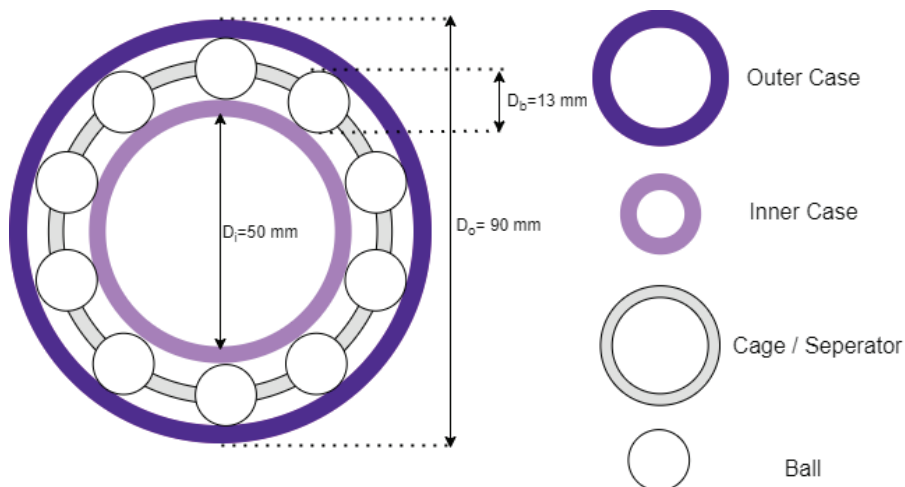$$D_i = 50$$
$$D_o = 90$$
$$D_b = 13$$
$$D_c = 63$$

Figure 6. The geometry of ball bearing.

509

Table 1. Theoretical bearing fault fundamental frequencies

Outer race fault fundamental frequency (Hz), FBo	Inner race fault fundamental frequency (Hz), FBi	Ball pass fundamental defect frequency (Hz), FBb	Cage/ Train fault fundamental frequency (Hz), FBc
99.2063	150.7937	57.9976	9.9206

Table 1 shows the calculated bearing fundamental fault frequencies. It is observed that the inner race defect excites with a high frequency among other faults, i.e., 150.8 Hz, while the cage defect excites with a low frequency among other faults, i.e., 10 Hz.

5.2 Fault Frequency Bands

Practical fault frequencies can differ from theoretical ones due to factors like load variations, RPM shifts, and operational changes. Additionally, when describing bearing component geometry, the thickness and contact angle may be approximated due to uncertainty. To tackle this challenge and improve fault frequency identification, a specific frequency band is chosen for each fault frequency, offering robustness against real-world variations. These bands, denoted as $[F-W/2]$, $[F+W/2]$ are cantered at the fundamental fault frequencies, f_F for inner case, outer case, bearing ball and cage/separator. Here, W represents the harmonic width, set at 0.55 times the shaft frequency to identify the required fault frequency peaks. In this work, the first two harmonics are considered.

So, for harmonic width W,

$$\frac{W}{2} = \frac{0.55x \ shaft \ frequency}{2}$$

$$\frac{W}{2} = \frac{0.55x25}{2} = 6.8750 \ Hz \tag{16}$$

For a given harmonic number, h, the signal bands are calculated from $f_1 = hf_F - \frac{W}{2}$ and $f_2 = hf_F + \frac{W}{2}$.

Within each frequency band, the highest value is determined, likely indicating the bearing fault frequency. Table 2 displays these frequency bands for four types of faults using fundamental and second harmonic frequencies. The first column (FB) contains of $F-W/2$ values, while the second holds $F+W/2$ values. FBc, FBo, FBi, and FBb represent the fault frequency bands for the cage, outer race, inner race, and ball defect. Features F69-F104 from these bands (Table 3) are combined with F1-F35 from Annexure I for classification.

Table 2. Fault frequency band

Fault band	Harmonics No.	$f_1 = hf_F - \frac{W}{2}$	$f_2 = hf_F + \frac{W}{2}$
FBc	I	3.0456	16.7956
	II	12.9663	26.7163
FBo	I	92.3313	106.0813
	II	91.5377	205.2877
FBi	I	43.9187	157.6687
	II	94.7123	308.4623
FBb	I	51.1226	64.8726
	II	109.1201	122.8701

Table 3. Features extracted from a diagonal slice of the bispectrum

| Feature No.	Feature	No. of features
F36	Sum of all peaks	1
F37-F52	Mean and max of all peak frequencies in the principal diagonal	16
F53-F68	Mean and max of all peak frequencies in d1	16
F69-F84	Mean and max of all peak frequencies in d2	16
F85-F100	Mean and max of all peak frequencies in d	16
F101	Sum of all peaks in the principal diagonal	1
F102	Sum of all peaks in d	1
F103	Sum of all peaks in d1	1
F104	Sum of all peaks in d2	1

6. FAULT CLASSIFICATION

This work explores supervised machine learning algorithms, which aims to predict data labels. The features are extracted from the Bispectrum as explained in section 5. The features matrix obtained from the Bispectrum of vibration data is of the size 2420x104. The data is split into 80:20 ratio. Where 80 % of the data is taken for training i.e. 2420 × 0.8= 1936 samples and the rest 20 % of the data which is 484 samples are used for testing. For the validation technique, a 10-fold cross-validation approach is chosen to train the training data due to its ability to mitigate overfitting. Subsequently, the classification process proceeds with the development of models. The current study employs various types of Decision Trees, Support Vector Machines, K-Nearest Neighbors, and Naive Bayes classifiers, each with distinct parameters such as splits, kernel functions, number of neighbors, and kernels, respectively. Sixteen classifiers (listed in Table 3) assess the performance of the extracted fault frequencies dataset. Machine learning techniques such as Decision Trees, Support Vector Machines, k-Nearest Neighbors, and Naive Bayes offer parameter adjustments beyond their defaults to fit dataset requirements. Each classifier has distinct optimizing parameters. For the decision tree, factors such as the maximum number of splits and split criterion are crucial. In this study, the impact of varying the number of splits on the decision tree's performance was examined. For KNN, optimizing parameters include the number of neighbors, distance metric, and distance weight. Here, the distance metric was set to Euclidean, while variations in the number of splits were explored, with distance weight held constant. Naïve Bayes optimization involves the distribution name and kernel type, with the distribution type fixed at Gaussian. The study evaluated Naïve Bayes performance by varying kernel types while keeping distribution

Table 4. List of classifiers

Classifier	Kernels	Notations
Decision Tree (T)	No. of Trees	T100
		T20
		T4
		T3
Support Vector Machine (S)	Gaussian	Sg
	Polynomial	S1
		S2
		S3
K Nearest Neighbors (K)	No. of neighbors	K1
		K3
		K10
		K50
		K100
Naive Bayes (NB)	Gaussian/ Normal	NBN
	Uniform/ Box	NBB
	Epanechnikov	NBE

Table 5. Overall performance of the classification model

Classifiers	Training accuracy	Train time (Sec)	Testing accuracy	Test time (Sec)
DT-100	100	6.6928	100	0.1705
DT-20	100	4.2167	100	0.1656
DT-4	100	4.7431	100	0.1579
DT-3	100	0.7098	100	0.1593
SVM-g	99.94	1.7776	99.945	0.1694
SVM-P1	99.954	9.5932	99.9726	0.1726
SVM-P2	99.8858	9.9482	99.8904	0.1732
SVM-P3	99.9574	10.1702	99.9863	0.182
KNN-1	100	0.8832	100	0.1636
KNN-3	99.9962	0.854	100	0.1634
KNN-10	99.9898	0.8675	100	0.17
KNN-50	99.7916	0.8857	99.8493	0.1705
KNN-100	99.5441	0.8987	99.4795	0.1711
NB-N	99.6877	22.6496	99.6301	0.4753
NB-B	94.9484	35.2905	95.0959	0.4543
NB-E	96.0822	29.7189	96.3973	0.4516
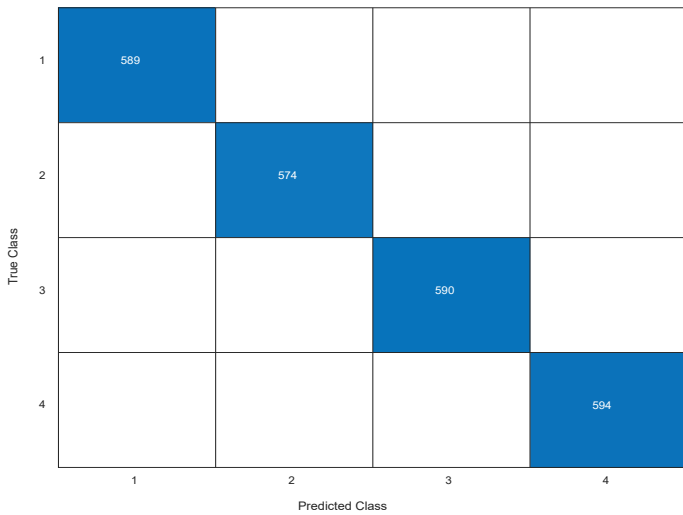
Decision Tree-3
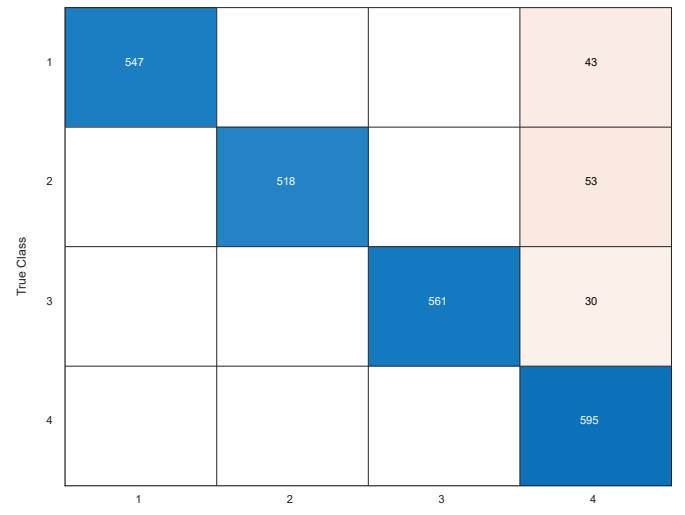
Naive Bayes − Gaussian/Box

Figure 7. Confusion matrix for the trained model.

type constant. SVM optimizing parameters encompass the kernel function, box constraint level, and kernel scale. Different kernels influence each classifier's behavior[21]. This research focused on varying kernel functions while keeping other parameters constant for assessing SVM performance. The varied parameters are listed in Table 4.

7. RESULTS AND DISCUSSION

The bispectrum of the bearing vibration data obtained from the vertical accelerometer is computed using the HOSA toolbox in MATLAB[22]. Features are extracted from the non-redundant region and bispectrum's diagonal slice for all faults. Scaling is primarily applied for distance-based algorithms like KNN and SVM, as they operate based on data point distances. Note that scaling occurs during data division, with central

parameters used during testing. Some classifiers, like decision tree-based algorithms, are insensitive to feature scaling since they split features independently. Sixteen classifiers are used, including Decision Tree (DT) with different splits, Support Vector Machine (SVM) with various kernels, K Nearest Neighbors (KNN) with different neighbors, and Naive Bayes (NB) with different kernel functions. Results are tabulated in Table 5.

This table displays mean accuracy, training, testing, and overall execution times across 100 Monte Carlo iterations. Monto Carlo is a phenomenon where classification is done for 100 different random possibilities. Further, in each iteration, a 10-fold cross-validation technique[23] is employed, wherein the dataset is split into ten equal folds. One fold serves as the validation set, while the remaining nine are used for training.
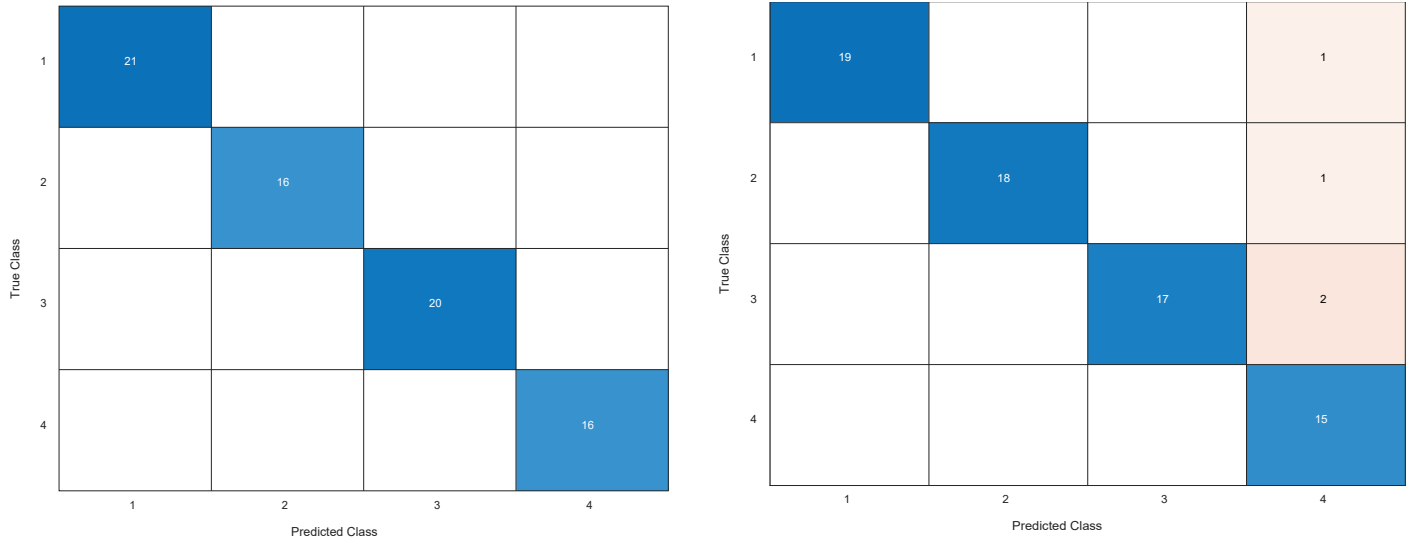
Decision Tree-3

Naive Bayes – Gaussian/ Box

Figure 8. Confusion matrix for tested model.

Figure 9. Comparison of training time and accuracy of machine learning algorithms with CNN.

This process repeats for each of the 100 Monte Carlo iterations, resulting in ten model trainings for each iteration.

Table 5 highlights that Decision Tree DT with all splits and K Nearest Neighbours with one neighbour achieve 100 % accuracy and quicker execution time (around 0.8 secs and 0.17 secs) in terms of training and testing respectively compared to other classifiers. This is attributed to their fewer splits (DT-3) and neighbors (KNN-1). In contrast, the Support Vector Classifier with polynomial kernel P3 offers good accuracy but takes longer to train than DT and KNN. Notably, Naive Bayes exhibits the longest training time among all classifiers. So for this data either DT or KNN would be an ideal classifier.

Figure 7 presents confusion matrices for DT-3 (highest accuracy) and Naive Bayes - Gaussian/Box (lowest accuracy). They are 4x4 matrices due to the four labels used. Additionally, Fig. 8 depicts a testing confusion matrix for one of the 100 Monte Carlo iterations with 484 data points subjected to a pre-trained model. Naive Bayes shows more misclassified data compared to other classifiers.

8. COMPARISON OF RESULTS

The results obtained in this paper are compared with our prior research[24], where a Convolutional Neural Network (CNN) was utilized for fault classification on the same dataset. It is noted that employing machine learning algorithms for fault classification requires less training time compared to CNN. This is understandable as CNN processes inputs as images and subsequently extracts features, which is time-consuming. In this instance, CNN took 6 minutes and 14 sec for training, whereas all the machine learning models required less time to achieve approximately the same level of accuracy. Decision tree and KNN-1 achieved 100 % accuracy with a maximum training time of 0.11 min, whereas CNN took considerably longer. This indicates that machine learning algorithms exhibited faster training compared to CNN. A comparison of the training times and accuracy are presented in Fig. 9.

9. CONCLUSION

In summary, this paper presents an innovative approach that

leverages Bispectrum analysis for the classification of bearing faults. The diagonal slice, a simpler and computationally more efficient alternative to the full bispectrum, is utilized, offering the advantage of containing valuable phase information for handling nonlinear systems. The study involves the extraction of a total of 104 features from the non-redundant region and frequency bands of the diagonal slice, followed by the evaluation of 16 supervised machine learning algorithms for their performance in bearing fault classification. Results indicate that the decision tree exhibited the highest accuracy, achieving a perfect score of 100 %, followed by KNN and SVM, while Naïve Bayes performed the least with an accuracy of 99.68 %. Additionally, it's noteworthy that in comparison to the Convolutional Neural Network (CNN) approach, the presented methodology showed superior performance in terms of accuracy and training time.

While this method has demonstrated superior performance, it entails additional work in feature extraction, which can be viewed as a limitation. Thus, there exists a trade-off between employing machine learning and deep learning algorithms. Selecting the optimal approach depends on the specific characteristics of the problem at hand. Future research should investigate hybrid approaches that integrate machine learning with other techniques to address these limitations and potentially improve the overall effectiveness of bearing fault detection systems.

REFERENCES

1. Nembhard, A.D.; Sinha, J.K.; Pinkerton, A.J. & Elbhbah, K. Condition monitoring of rotating machines using vibration and bearing temperature measurements. *In* Dalpiaz, G., *et al.* Advances in Condition monitoring of machinery in non-stationary operations. Lecture Notes In Mechanical Engineering, 2014, Springer, Berlin, Heidelberg.
    doi:10.1007/978-3-642-39348-8_13
2. Li, Z.; Yan, X.; Yuan, C.; Zhao, J. & Peng, Z. Fault detection and diagnosis of a gearbox in marine propulsion systems using bispectrum analysis and artificial neural networks. *J. Mar. Sci. App*. 2011, **10**, 17–24.
    doi: 10.1007/s11804-011-1036-7.
3. Guo, J.; Shi, Z.; Li, H.; Zhen, D.; Gu, F. & Ball, A.D. Early fault diagnosis for planetary gearbox based wavelet packet energy and modulation signal bispectrum analysis. *Sens*., 2018, **18**(9), 2908.
    doi: 10.3390/s18092908.
4. Zhang, R.; Wang, K.C.; Wu, L.; & Gao, H. Novel fault diagnosis approach for rolling-element bearings based on bispectral analysis. *Sens*. *Mater*., 2022, **34**(2), 765–778.
    doi: 10.18494/SAM3636.
5. Grover, C. & Turk, N. A novel fault diagnostic system for rolling element bearings using deep transfer learning on bispectrum contour maps. *Eng. Sci. Technol.*, 2022, **31**, 101049.
    doi: 10.1016/j.jestch.2021.08.006.
6. Wang, G.; Gu, F.; Rehab, I.; Ball, A.; & Li, L. A sparse modulation signal bispectrum analysis method for rolling element bearing diagnosis. *Math. Probl. Eng.*, 2018,, 12p, Article ID 2954094.
    doi: 10.1155/2018/2954094.
7. Guo, F.; Shao, Y.; Hu, N.; Fazenda, B. & Ball, A. motor current signal analysis using a modified bispectrum for machine fault diagnosis. *In* ICCAS-SICE 2009 - ICROS-SICE Joint Conference 2009 ,pp. 4890–4895, .
8. Alwodai, A.; Wang, T.; Chen, Z.; Gu, F.; Cattley, R. & Ball, A. A study of motor bearing fault diagnosis using modulation signal bispectrum analysis of motor current signals. *J. Signal Process*., 2013, **4**(3), 72–79.
    doi: 10.4236/jsip.2013.43b013
9. Berraih, S.A.; Baakek, Y.N.E. & Debbal, S.M.E.A. Preliminary study in the analysis of the severity of cardiac pathologies using the higher-order spectra on the heartbeats signals. *Pol. Med.Phys. Eng.*, 2021, **27**(1), 73–85.
    doi: 10.2478/pjmpe-2021-0010.
10. Jin, J.; Liu, C.; Daly, I; Miao, Y.; Li, S.; Wang, X. & Cichocki A. Bispectrum-based channel selection for motor imagery-based brain-computer interfacing. *IEEE Trans Neural Syst Rehabil Eng.*, 2020, **28**(10):2153-2163.
    doi: 10.1109/TNSRE.2020.3020975
11. Hashempour, Z.; Agahi, H. & Mahmoodzadeh, A. A novel method for fault diagnosis in rolling bearings based on bispectrum signals and combined feature extraction algorithms. *Sig. Image Process.*, 2022, SIViP 16, 1043–1051.
    doi:10.1007/s11760-021-02053-7
12. Sharma, A.; Patra, G.K. & Naidu, V.P.S. Bispectral analysis and information fusion technique for bearing fault classification. *Meas. Sci. Technol*., 2024, **35**, 015124,
    doi:10.1088/1361-6501/acffe4
13. McCormick, A.C.; & Nandi, A.K. Bispectral and trispectral features for machine condition diagnosis. IEE Proceedings - Vision, Image and Signal Processing, 1999, 146(5), 229–234.
    doi: 10.1049/ip-vis:19990673.
14. Alfatih, M.S.S.; Leong, M.S.; & Lim, M.H. Application of bispectral analysis in vibration fault detection: A review. *In* Applied Mechanics and Materials, 2014, 575, 464–468.
    doi: 10.4028/www.scientific.net/AMM.575.464.
15. Yang, B.; Wang, M.; Zan, T.; Gao, X. & Gao, P. Application of bispectrum diagonal slice feature analysis in tool wear states monitoring. *Int. J. Adv. Manuf. Technol. Res. Square.*, 2021.
    doi: 10.21203/rs.3.rs-775113/v1.
16. Chittora, A. & Patil, H.A. Significance of higher-order spectral analysis in infant cry classification. *Circuits, Syst. Signal Process*, 2018, **37**, 232–254.
    doi: 10.1007/s00034-017-0544-3.
17. Das, B.; Talukdar, M.; Sarma, R. & Hazarika, S.M. Multiple feature extraction of electroencephalograph signal for motor imagery classification through bispectral analysis. *Procedia Comput. Sci.,* 2016, **84**, 192–197.
    doi: 10.1016/j.procs.2016.04.086.
18. Mishra, M.; Pratiher, S; Banerjee, S. & Mukherjee, A. Grading heart sounds through variational mode decomposition and higher order spectral features. *IEEE*

Int. Instrum. Meas. Technol. Conf. Discov. (I2MTC), 2018, 1–5.
doi: 10.1109/I2MTC.2018.8409620.

19. Ghosh, T.; Biswas, T. & Khatun, R. A feature extraction scheme to classify motor imagery movements based on Bi-spectrum snalysis of EEG. *IOSR J. VLSI Signal Process.*, 2016, **6**(15), 28–35.
doi: 10.9790/4200-0605022835.

20. Wen Yean, C.; Wan Ahmad, W.K.; Mustafa, W.A.; Murugappan, M.; Rajamanickam, Y.; Adom, A.H.; Omar, M.I.; Zheng, B.S.; Junoh, A.K.; Razlan, Z.M.; et al. An emotion assessment of stroke patients by using bispectrum features of EEG Signals. *Brain Sci*. 2020, **10**, 672.
doi: 10.3390/brainsci10100672.

21. Kumar, P. & Hati, A.S. Review on machine learning algorithm based fault detection in induction motors. *Arch. Computat. Methods Eng.*, 2021, **28**, 1929–1940.
doi: 10.1007/s11831-020-09446-w

22. https://labcit.ligo.caltech.edu/~rana/mat/HOSA/, (Accessed on 02 February 2024).

23. Priva, T.H.; Shah, B.J.; Kulkarni, S.S. & Naidu, V. Bearing health condition monitoring using time-domain acoustic signal features. *In* IEEE 2nd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET), 2021, pp. 1-6.
doi: 10.1109/TEMSMET53515.2021.9768779

24. Bollineni, J.; Sharma, A. & Naidu, V. Bispectrum and convolution neural network based bearing fault diagnosis. *In* 4th International Conference on Circuits, Control, Communication and Computing (I4C), 2022, pp. 456-461.
doi: 10.1109/I4C57141.2022.10057647

CONTRIBUTORS

Ms Anju Sharma obtained her MTech degree in Aerospace Engineering from IIT Bombay. She is working as Principal Scientist at CSIR-NAL, Bangalore. She is currently pursuing her PhD from Academic council of Scientific and Innovative Research (AcSIR). Her research interest are Multisensor data fusion and condition monitoring and application of AI in condition monitoring.
In the current work she carried out the literature survey for the similar works, preparation of experimental setup, data acquisition, testing and conceptualised the study, and developed the code and manuscript preparation.

Dr G.K. Patra is working as a Chief Scientist in the Data Science and Supercomputing Division at the CSIR Fourth Paradigm Institute and is also a Professor at AcSIR, in the faculty of Mathematical and Information Sciences. His fields of interest include: Applications of AI/ML/DL to current generation Cyber physical systems and high performance computing.
In the current study, he provided guidance for the use of AI/ML tools and reviewed the manuscript.

Dr V.P.S. Naidu obtained ME in Medical Electronics from Gundy College of Engineering, Anna University Chennai and PhD from University of Mysore. He is working at Multi sensor data fusion lab, CSIR - NAL, Bangalore as a Chief Scientist and is also a Professor at AcSIR. His research interests are: Multi sensor data fusion, digital signal & image processing, condition monitoring and enhanced flight vision system.
In the current study, he provided guidance in experimental setup, testing and analysis of the results.

Annexure I
Bispectrum features from the Non-redundant region[9-10,17-20]

S. No	Feature name	Equation
F1	Normalised Bispectral entropy BE1	$BE1 = -\sum_n p_n \log p_n$ where $p_n = [B(f_1, f_2)] / \sum_\Omega |B(f_1, f_2)|$
F2	Normalised Bispectral squared entropy BE2	$BE2 = -\sum_n q_n \log q_n$ where $q_n = [B(f_1, f_2)]^2 / \sum_\Omega |B(f_1, f_2)|^2$
F3	Standard deviation of Bispectrum	2D standard deviation
F4	Mean magnitude of Bispectrum (MMOB)	$MMOB = \dfrac{1}{L} \sum_\Omega |B(f_1, f_2)|$ where L is the no. of points in the region Ω
F5	Weighted Center of Bispectrum WCOB for f1 frequency	$WCOB = \dfrac{\sum_\Omega i B(i,i)}{\sum_\Omega B(i,i)}$
F6	Weighted Center of Bispectrum WCOB for f2 frequency	$WCOB = \dfrac{\sum_\Omega j B(j,j)}{\sum_\Omega B(j,j)}$ Note: i,j denote the frequency bin index
F7	Sum of logarithmic amplitudes of Bispectrum (SOLA)	$SOLA = \sum_\Omega \log(B|f_1, f_2|)$
F8	Sum of logarithmic amplitudes of diagonal elements d1 (SOLADE)	$SOLADE(d1) = \sum_\Omega \log(B|d_1, d_1|)$
F9	Sum of logarithmic amplitudes of diagonal elements d2 (SOLADE)	$SOLADE(d2) = \sum_\Omega \log(B|d_2, d_2|)$
F10	Sum of logarithmic amplitudes of height elements d3 (SOLADE)	$SOLADE(d3) = \sum_\Omega \log(B|d_3, d_3|)$
F11	First order Spectral Moment of d1(FOSM)	$FOSM(d1) = \sum_{k=1}^{N} k \log |B(d_1, d_1)|$
F12	First order Spectral Moment of d2(FOSM)	$FOSM(d2) = \sum_{k=1}^{N} k \log |B(d_2, d_2)|$
F13	First order Spectral Moment of d3(FOSM)	$FOSM(d3) = \sum_{k=1}^{N} k \log |B(d_3, d_3)|$
F14	Second order Spectral Moment of d1 (SOSM)	$SOSM(d1) = \sum_{k=1}^{N} (k-FOSM)^2 \log |B(d_1, d_1)|$
F15	Second order Spectral Moment of d2 (SOSM)	$SOSM(d2) = \sum_{k=1}^{N} (k-FOSM)^2 \log |B(d_2, d_2)|$
F16	Second order Spectral Moment of d3 (SOSM)	$SOSM(d3) = \sum_{k=1}^{N} (k-FOSM)^2 \log |B(d_3, d_3)|$
F17	Sum of amplitudes of diagonal elements of d1	$Sum(d1) = \sum_{i=0}^{n} d1_i$

Sr. No	Feature name	Equation
F18	Sum of amplitudes of diagonal elements of d2	$Sum(d1) = \sum_{i=0}^{n} d2_i$
F19	Sum of amplitudes of diagonal elements of d3	$Sum(d3) = \sum_{i=0}^{n} d3_i$
F20	Simple Square Integral (SSI) of d1	$SSI(d1) = \sum_{i=0}^{n} \lvert d1_n^2 \rvert$
F21	Simple Square Integral (SSI) of d2	$SSI(d2) = \sum_{i=0}^{n} \lvert d2_n^2 \rvert$
F22	Simple Square Integral (SSI)of d3	$SSI(d3) = \sum_{i=0}^{n} \lvert d3_n^2 \rvert$
F23	Root mean square of d1 diagonal elements	$rms(d1) = \sqrt{\dfrac{1}{N} \sum_{n=1}^{N} d1_n^2}$
F24	Root mean square of d2 diagonal elements	$rms(d2) = \sqrt{\dfrac{1}{N} \sum_{n=1}^{N} d2_n^2}$
F25	Root mean square of d3 diagonal elements	$rms(d3) = \sqrt{\dfrac{1}{N} \sum_{n=1}^{N} d3_n^2}$
F26	Variance of d1	$\mathrm{var}(d1) = \dfrac{1}{N-1} \sum_{n=1}^{N} d1_n^2$
F27	Variance of d2	$\mathrm{var}(d2) = \dfrac{1}{N-1} \sum_{n=1}^{N} d2_n^2$
F28	Variance of d3	$\mathrm{var}(d3) = \dfrac{1}{N-1} \sum_{n=1}^{N} d3_n^2$
F29	V3 order of d1	$V3(d1) = \left(\dfrac{1}{N} \sum_{i=1}^{N} \lvert d1 \rvert_i^3 \right)^{1/3}$
F30	V3 order of d2	$V3(d2) = \left(\dfrac{1}{N} \sum_{i=1}^{N} \lvert d2 \rvert_i^3 \right)^{1/3}$
F31	V3 order of d3	$V3(d3) = \left(\dfrac{1}{N} \sum_{i=1}^{N} \lvert d3 \rvert_i^3 \right)^{1/3}$
F32	Log Detector of d1	$Log = e^{\frac{1}{N}\sum_{n=1}^{N} \log\lvert d1_n \rvert}$
F33	Log Detector of d2	$Log = e^{\frac{1}{N}\sum_{n=1}^{N} \log\lvert d2_n \rvert}$
F34	Log Detector of d3	$Log = e^{\frac{1}{N}\sum_{n=1}^{N} \log\lvert d3_n \rvert}$
F35	Bispectrum-magnitude variability	$A_{\mathrm{var}} = \dfrac{1}{L} \sum_{\Omega} \lvert B(f_1, f_2) \rvert - MMOB$