A Machine Learning Approach for Bearing Fault Identification Using IFWHT, GLCM, and Feature Ranking

UDC:621.822:004.8
Original scientific paper https://doi.org/10.46793/adeletters.2026.5.1.1

Vipul Dave[1*], Pradeep Kumar Karsh[1], Umangbhai Soni[2], Vipul M Dabhi[3], Khyati Zalawadia[4], Patel Ankit Rameshbhai[5]

[1]Department of Mechanical Engineering, Parul Institute of Engineering & Technology, Parul University, 391760, Vadodara, Gujarat, India
[2]Department of Electronics & Communication Engineering, Parul Institute of Engineering & Technology, Parul University, 391760, Vadodara, Gujarat, India
[3]Department of Computer Science Engineering, Gokul Global University, Sidhpur, 384151, Gujarat, India
[4]Department of Computer Engineering, Parul Institute of Engineering & Technology, Parul University, 391760, Vadodara, Gujarat, India
[5]Department of Aeronautical Engineering, Parul Institute of Engineering & Technology, Parul University, 391760, Vadodara, India

Abstract:
Time–frequency analysis is necessary for condition monitoring of rotating machinery; comparatively limited attention has been given to image-based texture feature analysis derived from time–frequency representations. A bearing fault diagnosis model is proposed that combines the Inverse Fast Walsh-Hadamard Transform (IFWHT), texture feature extraction, feature ranking, and machine learning classification. First, IFWHT transforms vibration signals into time-frequency images, thereby improving the effective representation of transient and fault-related data. Based on these images, two-dimensional texture attributes are derived from the Gray Level Co-occurrence Matrix (GLCM), which records the spatial relationships and structural patterns of various bearing health conditions. The ReliefF algorithm ranks and selects the most discriminative features, reducing feature redundancy and improving classification performance. Machine learning classifiers are then trained and evaluated using the ranked feature set. The findings show that the suggested ranked feature-based framework provides consistent and good-quality fault classification performance. Fine KNN and Cubic SVM are among the compared models, where the former will reach a maximum classification accuracy of 93.8%, and the latter will reach an accuracy of 91.7% at various combinations of faults. The methodology suggested offers a strong and adaptable approach in the context of fault diagnosis, which effectively promotes predictive maintenance by lessening unforeseen downtimes and increasing the dependability of operations in industries.

ARTICLE HISTORY
Received: 13 September 2025
Revised: 19 January 2026
Accepted: 31 January 2026
Published:31 March 2026

KEYWORDS
Bearing Faults, IFWHT, GLCM, ReliefF, Ten-fold, CWRU, SVM, KNN, Reliability

1. INTRODUCTION

Rolling element bearings are vital components in industrial machinery, and their failure can result in equipment downtime and substantial production losses. Therefore, early and accurate detection of bearing faults is critical to avoid unexpected failures and associated operational disruptions. Over recent

*CONTACT: Vipul Dave, e-mail: vipul.dave270188@paruluniversity.ac.in

decades, numerous diagnostic methodologies have been proposed for fault detection in bearings under both operational and non-operational states [1]. In addition to vibration-based fault diagnosis, temperature monitoring techniques have shown that measured temperature profiles correlate with internal friction conditions and can effectively reflect bearing health. Among these, vibration-based diagnostics have gained increased importance due to their high efficiency in monitoring machine health [2]. Still, one of the primary challenges in vibration-based scrutiny lies in the use of the inherently non-stationary nature of vibration signals, which always changes over time. To manage such complications, signal processing techniques in both time and frequency domains have been explored [3]. Although the Fast Fourier Transform (FFT) is commonly applied in the detection of early bearing faults, the detection capability of this method is limited by hiding weak fault-related signals in the presence of background noise. The limitation lowers its sensitivity, mainly in the case of non-stationary vibration signals, which are common in rotating machinery. As one way of overcoming these issues, higher time-frequency domain methods, including Short-Time Fourier Transform (STFT), Wavelet Transform, and Empirical Mode Decomposition (EMD), have been mostly used in bearing fault diagnosis [4]. All these techniques have their own benefits, but the performance of each highly depends on the types of signals and diagnostic parameters. It is important to note that EMD has some natural disadvantages, like mode mixing and the extraction of the intrinsic mode function with large frequency bands and overlapping frequency bands, which reduce the clarity of the fault features. To address these issues, the inverse Fast Walsh-Hadamard Transform (IFWHT) has become an important focus because this tool is efficient in fault diagnosis [5]. IFWHT is an appropriate method because it is computationally efficient in large datasets and offers a high-resolution time-frequency representation of non-stationary vibration signals. Also has a higher ability to maintain fault-induced impulsive features than the traditional Fourier-based methods. The success of IFWHT is based on the fact that it is effective in the extraction of relevant fault-related information on complex vibration signals. Overcoming the drawbacks of FFT, STFT, and EMD, IFWHT is highly applicable in real-life industry fault diagnosis [6]. In the context of bearing faults, various types of defects, such as Ball defect (BD), Outer race defect (ORD), Inner race defect (IRD), and cage defect, exhibit different characteristics in vibration signals due to operating conditions, geometry, and friction at surfaces. To effectively analyze and diagnose these faults, it is necessary to represent fault characteristics in 2D images [7].

The use of texture features based on the time-frequency representations has shown to be useful in improving the precision and dependability of a fault diagnosis system. Two of the most noticeable types of texture analysis techniques used are the Gray-Level Co-occurrence Matrix (GLCM) and Local Binary Patterns (LBP), which have demonstrated significant success in a variety of applications in the engineering, diagnostic, and biomedical imaging world. Particularly, GLCM has been found to be very useful in capturing challenging spatial relations in the texture features obtained through the time-frequency images, and hence it may be a strong tool in the fault diagnosis bearings. Furthermore, the latest research has proposed methodologies that affect adaptive optimal kernel time-frequency representations along with uniform LBP fault ID to provide good diagnostic performance [8].

In an attempt to further increase diagnostic accuracy, a variety of feature selection methods have been investigated, which try to keep only the most informative features and eliminate the redundant or uninformative ones. The methods of Information Gain, ReliefF, Fisher Score, and Mahalanobis Distance are fairly popular among feature ranking tasks [9]. In the present research, the authors present an elaborate method that leads to the extraction of two-dimensional texture characteristics on the images of Fast Walsh-Hadamard Transform (FWHT)-based time-frequency. The ReliefF algorithm is used in ranking and selecting the most important features so that it can then be used to build a fine feature vector.

The proposed methodology is verified using Case Western Reserve University (CWRU) data, and the ten-fold cross-validation method is used to achieve the necessary statistical accuracy and exclude bias. The combination of enhanced functions and machine learning classifications, including Support Vector Machines (SVMs), KNN, and others, results in significantly higher fault classification rates.

2. MATERIALS AND METHODS

Fig. 1 represents the organization of the present work; the signals of vibrations are initially obtained during the test rig (CWRU) on the bearing in various conditions of operation. The entire signal set is subsequently subjected to the IFWHT process to improve fault-related information. The treated signals are processed to produce spectrograms that

give time-frequency representations. The spectrograms are used to generate relevant features and rank them using ReliefF feature selection. Lastly, the applied features are working to train an ML classifier with a tenfold cross-validation, to identify the bearing fault correctly.

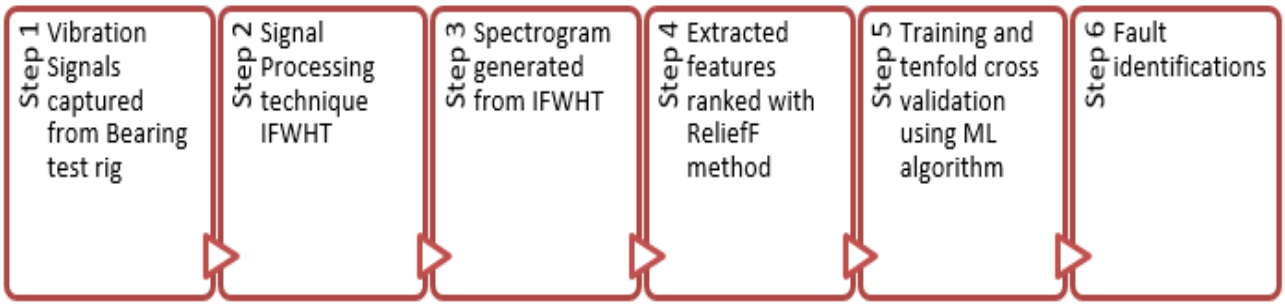


Fig. 1. Flow chart for Fault diagnosis

2.1 Inverse Fast Walsh Hadamard Transform (IFWHT)

The Inverse Fast Walsh-Hadamard Transform (IFWHT) is a fast, efficient mathematical operation that is applied to restore the original signal based on its Walsh-Hadamard domain representation. It is the reverse action of the Fast Walsh-Hadamard Transform (FWHT) that is commonly known in terms of small computational cost and commercial viability in high-speed signal processing [10]. In contrast to Fourier-based transforms, the Walsh-Hadamard Transform uses orthogonal square-wave basis functions and is thus especially useful for signals with sharp edges and sparse properties. The IFWHT works through recursive interruption of the transformed signal into smaller subs-blocks and a sequence of structured sequences of additions and subtractions on the real values. This divide-and-conquer method also greatly simplifies the computational complexity (N) to (N), where N is the signal length. A major benefit of the IFWHT is that only real-valued arithmetic is required, so complex multiplication is not needed. This property is easy to implement and minimizes numeric errors, and this is more advantageous in real-life engineering. The IFWHT is particularly useful with signals in which sharp discontinuities, impulsive content, or sparse data occur [11]. These features are usually found in mechanical vibration signals, fault transients, and condition monitoring signals. The replication of such signals is precisely done by the IFWHT so that the features that are important to the diagnosis are not distorted or lost as they would otherwise be with the conventional transforms.

The IFWHT is used in signal processing and image reconstruction, denoising, and feature enhancement. It is useful in data compression to allow efficient and loss-free representation of signals, allowing less storage space to be used without information loss. Moreover, the IFWHT can be used in pattern recognition and feature extraction, in which transformed coefficients reveal discriminative signal properties. Recently, feature sets based on IFWHT have achieved attainment in being combined with machine learning algorithms to solve classification problems, such as fault type and condition assessment [12]. The reconstructed and transformed signals give reduced representations that are compact and reduce noise effectively to enhance the performance of classifiers. Because of the rapid calculation, ease, and ability to deal with non-stationary and discontinuous signals, the IFWHT is a stable and efficient method to be used in the complex analysis of signals and smart fault detection systems.

2.2 Gray Level Co-Occurrence Matrix (GLCM)

The end result is time-recurrence images that were obtained after the Inverse Fast Walsh Hadamard Transform (IFWHT) was used. Overall, meaningful information about the working condition of a bearing cannot be obtained directly out of raw image-based representations since the former is mostly composed of surface texture and tonal content with no clear fault-related features. A Gray Level Co-Occurrence Matrix (GLCM) can be applied to derive this data [13]. GLCM is a popular statistical method used to extract textural features of images, which is a histogram of pixels that are separated by constant spatial relationships. The GLCM is applied to a digital image, which is represented as a two-dimensional array of resolution cells, and measures the frequency with which a pixel of a given intensity value succeeds in a given spatial relationship to another pixel of a given value [14]. The GLCM function counts the

occurrence of various combinations of pixels and records this information by using the various elements of the matrix. The images can be analyze to extract the effective textural features of the images, which is useful in the analysis of bearing working states and other uses [15]. Pixel distance (d): 1 pixel, Orientation angles (θ): 0°, 45°, 90°, and 135°, Gray levels: 256, the final GLCM feature values were obtained by averaging the feature responses over all four orientations, ensuring rotational robustness. Fig. 2 shows GLCM Images at 1797 rpm for fault size (0.17mm, 0.35mm, 0.53mm).

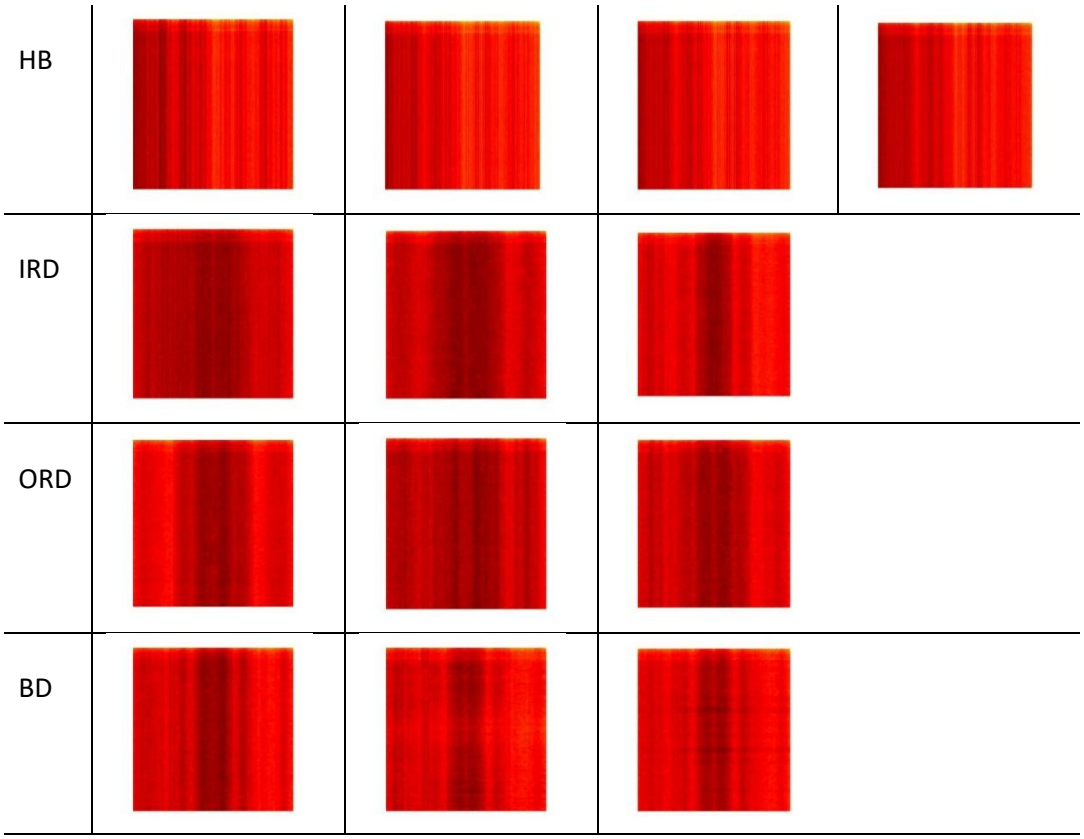


Fig. 2. GLCM Images at 1797 rpm for fault size (0.17 mm, 0.35 mm, 0.53 mm)

2.3 Feature Selection Using ReliefF

ReliefF is an effective and popular feature selection technique in machine learning and data mining. It addresses the original Relief algorithm, which is a classification task, to solve feature selection problems. ReliefF is mainly aimed at giving weights to every attribute in a dataset, which indicates their significance in classification or regression problems [16]. The weight assignment is grounded in the examination of the distances between samples in the feature space, with the help of the information about nearest neighbors. The algorithm follows a number of major steps. For every instance in the dataset, ReliefF finds its nearest neighbors from both the same class (positive instances) and different classes (negative instances). It then calculates the feature-value differences between the current instance and its nearest neighbors from both the positive and negative classes. These variances offer insight into the features' relevance in discriminating between classes. ReliefF then updates the feature weights by arranging those with more significant differences in values between positive and negative neighbors [17]. This process is repeated for each instance, and the final feature weights are obtained by averaging across all instances.

The MATLAB implementation of the ReliefF algorithm with the default nearest neighbor value of k = 10, which represents a good exchange between sensitivity and robustness in multiclass tasks, was used to perform feature ranking. The remaining algorithmic parameters were kept as defaults in MATLAB. The Gray Level Co-occurrence Matrix (GLCM) possesses characteristic features of describing the spatial distribution, dependency, and complexity of texture patterns in time-frequency images generated by vibration signals. Primary measurements of texture uniformity and smoothness include energy, homogeneity, inverse difference, inverse difference normalized, and inverse difference moment normalized, which can effectively differentiate between faulty and healthy

bearing conditions. Local intensity variations and heterogeneity are represented by contrast, dissimilarity, difference variance, and sum variance, and are more intense when defects are severe because they cause impulsive vibration behavior. Correlation and auto-correlation are used to define a linear dependency and period between adjacent gray levels and, therefore, are useful in detecting repetitive fault-related patterns in the rotating machine. Entropy, sum entropy, and difference entropy indicate randomness and complexity in the texture structures, which generally tend to increase under defective operating states. Correlation statistics (IMC1 and IMC2) are measures of the linear and nonlinear gray-level dependencies that are more sensitive to smaller texture variations that can be related to initial fault phases. The sum average and sum of squares variance are global energy distribution and dispersing within the texture, which aids in characterizing fault intensity. Maximum probability indicates prevailing texture patterns associated with consistent vibration patterns. Cluster shade and cluster prominence represent the higher-order statistical moments, which represent the asymmetry and peaking of texture distributions that are strong in harsh faults. Together, these GLCM features offer a full representation of the spatial texture properties and, thus, they are very useful in bearing fault diagnosis, surface examination, pattern recognition, and condition monitoring applications. Two-dimensional features are outlined in Fig. 3.

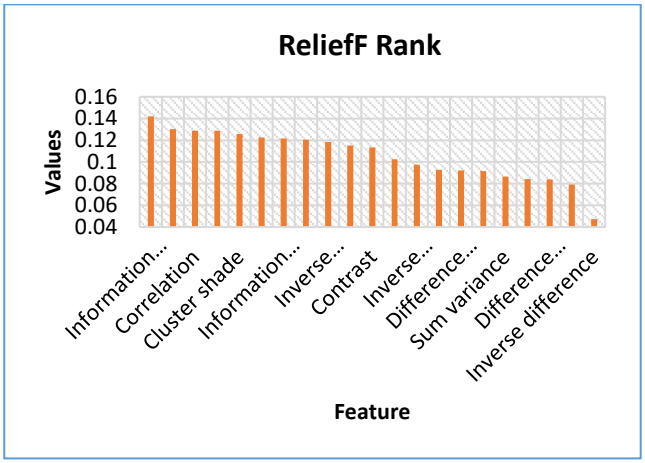


Fig. 3. ReliefF feature rank

2.4 Support Vector Machine

Support Vector Machine (SVM) is a highly effective supervised learning algorithm that has gained wide application across classification and regression tasks. The primary objective of SVM is to categorize an optimal hyperplane that specifically separates instances belonging to different classes within a high-dimensional feature space. This separation is achieved by maximizing the margin, the distance between the hyperplane and the nearest data points from each class, thereby enhancing the model's generalization capability and reducing overfitting risk [18]. A separate feature of SVM lies in its dependence only on support vectors, the critical training instances that lie closest to the decision boundary, thus snubbing the rest of the dataset during the decision-making process. This characteristic not only simplifies calculation but also strengthens the model's strength. Furthermore, by the use of kernel functions, SVM may convert non-linearly separable input data into a higher-dimensional data space where a linear separation is possible. These kernels (polynomial, radial basis function, or sigmoid) allow SVM to deal with non-linear and complicated decision boundaries effectively. With its scalability, high-dimensional data manipulation ability, and solid theoretical framework, SVM continues to be an excellent method when it comes to complex data structures and small training sets [19]. For SVM, the default parameters are:

- Kernel function: Linear;
- Kernel scale: Automatic;
- Box constraint (C): 1;
- Standardization: Default.

2.5 K nearest Neighbor (KNN)

The K-Nearest Neighbors (KNN) system is a non-parametric, instance-based learning method extensively known for its simplicity and usefulness in classification and regression problems. It works on the fundamental assumption that similar input samples tend to produce similar outputs, following the principle of locality in feature space. In a common KNN operation, a set of labeled samples is set, and the parameter k is set to define the number of close neighbors that will help predict a new, unlabeled sample [20]. The choice of k is important to strike a balance between model sensitivity and prediction stability. Choosing smaller values of k can cause the model to be sensitive to noise, whereas wider choices can eliminate local irregularities but can miss finer models. Upon receiving a new data point, KNN computes the distance (commonly using Euclidean or Manhattan metrics) between this point and all labeled instances in the training set. The algorithm then identifies the 'k' closest data points and, for classification, assigns the majority

class among these neighbors to the new input. For regression, the predicted value is typically the mean or weighted average of the target values of the nearest neighbors [21]. Despite its computational intensity in large datasets, KNN's versatility, ease of implementation, and strong empirical performance make it a valuable baseline and benchmarking tool in many practical machine learning applications. For KNN, the default parameters are:

- Number of neighbors (k): 1;
- Distance metric: Euclidean distance;
- Distance weighting: Equal (no weighting);
- Standardization: Disabled by default.

The choice of k = 1 allows fine local decision boundaries and is generally used as a baseline in pattern recognition tasks, especially when feature sets are well separated. Although the KNN classifier achieved the highest classification accuracy (up to 93.8%), it is known that instance-based learners may be inclined to overfitting, especially when the number of features increases or when noisy samples are present. However, in the present study, the use of ReliefF-based feature selection helped alleviate this risk by retaining only the most unfair features, thereby enhancing generalization performance.

3. EXPERIMENTAL PROCEDURE

Table 1 presents the geometric specifications of the deep-groove ball bearing used in the CWRU dataset.

Table 1. Bearing Specifications

Ball diameter (mm)	Inside diameter (mm)	Outside diameter (mm)	Rolling element (Z)	Contact angle (α)
7.94	25.01	51.99	9	0

Fig. 4 shows the Case Western Reserve University (CWRU) Bearing Data Center, a widely used laboratory database for bearing fault diagnosis and condition monitoring studies. Comprises a 2 HP (approx. 1.492 kW) motor, a torque transducer/encoder, a dynamometer, and control electronics. Test bearings supporting the motor shaft endured electro-discharge machining to introduce single-point faults with diameters ranging from approximately 0.18 to 1.02 mm. SKF deep groove ball bearings (SKF 6205-2RS JEM) were used to study fault sizes. Vibration data was collected using accelerometers (PCB Piezotronics 353B33 piezoelectric accelerometer) positioned at the 12 o'clock position on the motor housing. Data recorded at 12,000 samples per second for drive-end bearing faults. The recorded vibration signals were acquired from the drive-end bearing and correspond to distinct bearing conditions: healthy bearing (HB), inner race defect (IRD), ball defect (BD), and outer race defect (ORD). The data were collected at rotational speeds of 1730, 1750, 1772, and 1797 rpm, comprising four samples for HB, sixteen samples each for BD and IRD, and twenty-eight samples for ORD.

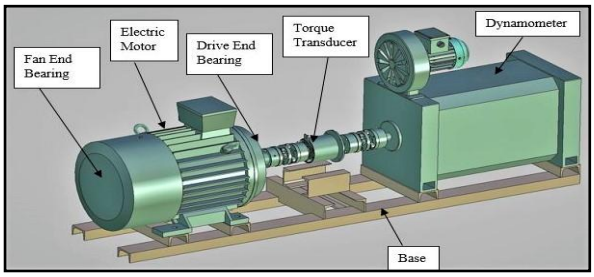


Fig. 4. Bearing test rig

The Inverse Fast Walsh Hadamard Transform is applicable for time-frequency-based image generation, followed by the calculation of two-dimensional features through GLCM. Visual representations of bearing conditions are presented in Fig. 2, each corresponding to specific states at a consistent speed of 1797 rpm for comparison purposes. Notable differences are presumed among surfaces depicting various fault conditions. Notably, the intensity of the black description increases as the bearing condition shifts from normal to defective, assisting improved fault detection across varying bearing states.

4. RESULTS AND DISCUSSION

The performance of the proposed IFWHT–GLCM–ReliefF-based bearing fault diagnosis structure was evaluated using Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Subspace KNN Ensemble classifiers. A ten-fold cross-validation approach was adopted to ensure healthy and neutral performance assessment. The vibration dataset consisted of four bearing conditions: Healthy Bearing (HB), Ball Defect (BD), Inner Race Defect (IRD), and Outer Race Defect (ORD). A total of 21 ranked GLCM texture features were used for classification.

4.1 Effect of Feature Ranking on Classification Performance

The ReliefF feature ranking was used to identify the most unfair texture features extracted from

IFWHT-based time–frequency images. The ranking results (Fig. 3) indicate that features related to contrast, correlation, entropy, homogeneity, and higher-order texture statistics contribute significantly to fault separation. Feature ranking effectively reduced redundancy and improved classifier generalization by retaining only the most relevant features.

4.2 Classification Results for Different Fault Combinations

Figs. 5–8 present the ten-fold cross-validation accuracy (%) for different fault class combinations, where the x-axis represents ranked feature indices and the y-axis denotes classification accuracy.

For the HB–BD–IRD classification (Fig. 5), Cubic SVM achieved a maximum accuracy of 91.7% using five selected features, demonstrating strong generalization with a compact feature set. Fine KNN also performed competitively, achieving 91.7% accuracy with seven features. The Ensemble classifier showed lower accuracy, suggesting limited generalization.

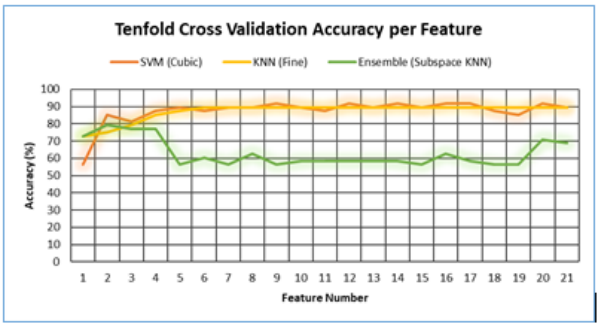


Fig. 5. Tenfold accuracy of HB_BD_IRD

In the HB–IRD–ORD arrangement (Fig. 6), Fine KNN achieved the highest accuracy of 93.8% with 19 selected features, highlighting its strong point in capturing local decision boundaries when sufficient discriminative features are available. Cubic SVM maintained stable performance with accuracies consistently above 87.5%, reflecting its robustness across varying feature subsets.

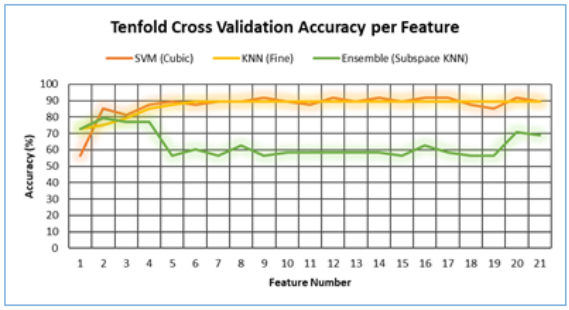


Fig. 6. Tenfold accuracy of HB_IRD_ORD

For the HB–BD–ORD classification (Fig. 7), Fine KNN demonstrated consistent performance with an average accuracy of 89.6% across multiple feature sets. Cubic SVM achieved slightly lower but stable accuracy, while the Ensemble method showed clear signs of overfitting, achieving high training accuracy but poor cross-validation performance.

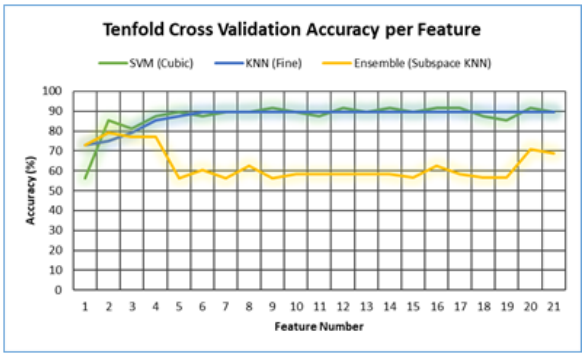


Fig. 7. Tenfold accuracy of HB_BD_ORD

In the most challenging HB–BD–IRD–ORD multi-class classification case (Fig. 8), Fine KNN again outperformed the other classifiers, achieving a maximum ten-fold accuracy of 81.3% using features ranked between 16 and 21. Cubic SVM achieved moderate accuracy (up to 82.8%), whereas the Ensemble classifier performed the weakest, supporting its unsuitability for this dataset.

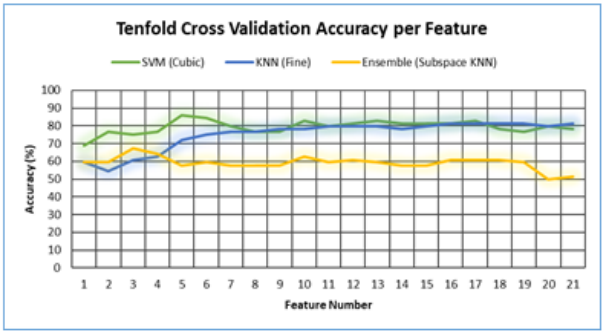


Fig. 8. Tenfold accuracy of HB_BD_IRD_ORD

4.3 Comparative Analysis of Classifiers

Table 2 summarizes the classification performance across all bearing conditions. The results indicate that:

- Fine KNN achieves the highest peak accuracy (93.8%), particularly in the HB–IRD–ORD condition, due to its ability to use local feature distributions.
- Cubic SVM provides more stable and consistent performance with fewer features, achieving up to 91.7% accuracy while maintaining better generalization.

• Subspace KNN Ensemble consistently underperforms, showing overfitting and lower cross-validation accuracy.

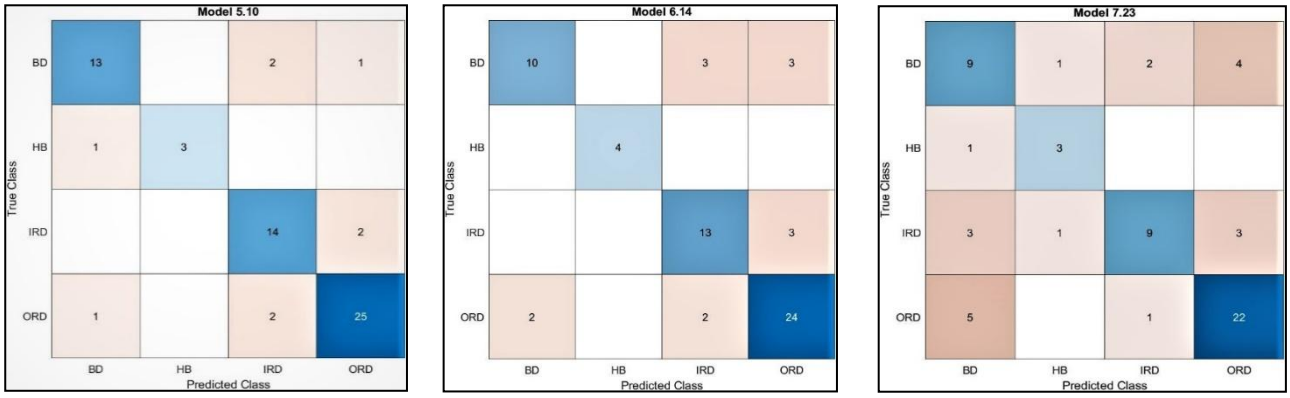
Table 2. Comparison study between different machine learning classifiers

Classifier	Bearing Condition	Number of Selected Features	Tenfold Accuracy (%)
SVM (Cubic)	HB–BD–IRD	5	91.7
	HB–BD–ORD	9	91.7
	HB–IRD–ORD	10	89.6
	HB–BD–IRD–ORD	5	85.9
KNN (Fine)	HB–BD–IRD	7	91.7
	HB–BD–ORD	6	89.6
	HB–IRD–ORD	19	93.8
	HB–BD–IRD–ORD	16	81.3
Ensemble (Subspace KNN)	HB–BD–IRD	10	83.3
	HB–BD–ORD	2	79.2
	HB–IRD–ORD	3	81.3
	HB–BD–IRD–ORD	3	68.8

These observations confirm that classifier performance is highly dependent on feature quality and fault complexity. While KNN excels in peak accuracy, SVM remains a reliable choice when stability and reduced feature dimensionality are required.

4.4 Confusion Matrix Analysis

The confusion matrices shown in Fig. 9 and Table 3 provide class-wise performance insights. The Cubic SVM classifier demonstrates superior discrimination for IRD and ORD faults, with minimal misclassification. Fine KNN performs strongly across all fault categories, particularly for HB and BD conditions, though minor confusion is observed between BD and IRD. In contrast, the Ensemble classifier exhibits significant misclassification across fault classes, especially between BD and ORD, further validating its limited effectiveness.



(a) SVM Cubic (b) Fine KNN (c) Ensemble (subspace KNN)

Fig. 9. Confusion matrix for all classifiers at the 5th feature

Table 3. Confusion matrix

SVM Cubic	**BD**	**13**	0	2	1	**BD**	**14**	0	2	**BD**	**12**	0	4
	HB	1	**3**	0	0	**HB**	1	**3**	0	**HB**	0	**4**	0
	IRD	0	0	**14**	2	**IRD**	0	0	**16**	**ORD**	0	0	**28**
	ORD	1	0	2	**25**		**BD**	**HB**	**IRD**		**BD**	**HB**	**ORD**
		BD	**HB**	**IRD**	**ORD**								
KNN (Fine KNN)	**BD**	**10**	0	3	3	**BD**	**15**	0	1	**BD**	**13**	0	3
	HB	0	**4**	0	0	**HB**	0	**4**	0	**HB**	0	**4**	0
	IRD	0	0	**13**	3	**IRD**	2	0	**14**	**ORD**	2	0	**26**
	ORD	2	0	2	**24**		**BD**	**HB**	**IRD**		**BD**	**HB**	**ORD**
		BD	**HB**	**IRD**	**ORD**								
Ensemble (subspace KNN)	**BD**	**9**	1	2	4	**BD**	**15**	0	1	**HB**	**14**	0	2
	HB	1	**3**	0	0	**HB**	0	**4**	0	**BD**	1	**3**	0
	IRD	3	1	**9**	3	**IRD**	5	0	**11**	**ORD**	7	0	**21**
	ORD	5	0	1	**22**		**BD**	**HB**	**IRD**		**HB**	**BD**	**ORD**
		BD	**HB**	**IRD**	**ORD**								

4.5 Discussion and Comparison with Existing Studies

The comparative study in Table 4 indicates that the proposed method achieves competitive performance relative to existing bearing fault diagnosis approaches. Although the deep learning-based approaches (like CNNs) have been reported as having greater accuracy, they are more expensive to implement in large datasets. By comparison, the proposed IFWHT-GLCM-ReliefF model is more accurate but has a smaller complexity, and thus this model is applicable in real-time and industrial condition monitoring. Furthermore, consistent with observations reported by Ul Haq et al. (2025), classifier effectiveness is application-and feature-dependent [22]. Similar to their findings, ensemble methods do not necessarily ensure superior performance unless supported by well-structured and diverse feature representations.

Table 4. Comparison of present work with previous published work

Authors	Model	Dataset	Fault Types	Accuracy (%)
Hoang and Kang [23]	CNN	CWRU	NS, IRF, ORF, BF	97.74
Kavathekar et al. [24]	Rotation Forest	CWRU	NS, ORF, IRF, BF	75.00
Konar and Chattopadhyay [25]	CWT	Experimental	NS, FS	73.33, 90.00, 93.33, 96.67, 100.00
Kankar et al. [26]	CWT	Experimental	NS, IRF, ORF, BF	76.00, 94.66, 98.66
Li et al. [27]	FLC	CWRU	NS, IRF, ORF, BF	92.00, 96.00, 100.00
Ul Haq et al. [22]	Random Forest	Network intrusion detection systems (NIDS) dataset	-NA-	99.78
	SVM			53.14
	AdaBoost			99.08
	KNN			98.49
	Gaussian Naive Bayes			55.72
	Multinomial Naive Bayes			55.14
This Study	SVM (Cubic)	CWRU	HB, BD, IRD	91.70
	Fine KNN		HB, IRD, ORD	93.80
	Subspace KNN		HB, BD, IRD	83.30

5. CONCLUSION

This research presents an efficient bearing fault diagnostics framework that uses an inverse Fast Walsh-Hadamard Transform-based time-frequency representation, along with texture features and machine learning classification. The proposed method enables improved visualization and characterization of fault patterns under varying bearing conditions by transforming non-stationary vibration signals into informative two-dimensional time-frequency images.

In order to enhance the reliability of classifications, the ReliefF algorithm was used to rank and select the most discriminative features, hence decreasing the redundancy of features and enhancing the performance. Support Vector machine and K-Nearest neighbors classifiers under tenfold cross-validation were used to validate the effectiveness of the ranked feature set. The outcome of the experiment on CWRU bearing data reveals that the systematic framework delivers high diagnostic reliability with Fine KNN and Cubic SVM, demonstrating stable, consistent performance across various fault combinations.

The comparison shows that the proposed method strikes a balance between accurate fault diagnosis and reasonable computational effort, compared with existing methods. In contrast to deep learning models that require large amounts of training data and high computational speed, the existing outline is small and interpretable to be included in realistic applications of condition monitoring.

Future research can best be expected at expanding the suggested methodology to other

rotating machinery parts (gears and turbines), as well as additional time-frequency representation and the incorporation of better ensemble/deep learning models to increase the diagnostic capacity of the method in more complex operating conditions.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] W. Song, W. Shen, L. Gao, X. Li, An early fault detection method of rotating machines based on unsupervised sequence segmentation convolutional neural network. *IEEE Transactions on Instrumentation and Measurement*, 71, 2021: 1-12. https://doi.org/10.1109/TIM.2021.3132989

[2] E. Desnica, A. Ašonja, M. Kljajin, H. Glavaš, A. Pastukhov. Analysis of Bearing assemblies refit in agricultural PTO shafts. *Tehnički Vjesnik*, 30(3), 2023: 872-881. https://doi.org/10.17559/TV-20221117162133

[3] A. Asonja, E. Desnica, Reliability of agriculture universal joint shafts based on temperature measuring in universal joint bearing assemblies. *Spanish Journal of Agricultural Research*, 13(1), 2015: e0201-e0201. https://doi.org/10.5424/sjar/2015131-6371

[4] T. Govardhan, A. Choudhury, Fault diagnosis of dynamically loaded bearing with localized defect based on defect-induced excitation. *Journal of Failure Analysis and Prevention*, 19, 2019: 844-857. https://doi.org/10.1007/s11668-019-00668-0

[5] B. Zhang, W. Wang, Y. He, A hybrid approach combining deep learning and signal processing for bearing fault diagnosis under imbalanced samples and multiple operating conditions. Scientific Reports, 15, 2025: 13606. https://doi.org/10.1038/s41598-025-98138-1

[6] D. Wu, J. Wang, H. Wang, H. Liu, L. Lai, T. He, T. Xie, An Automatic Bearing Fault Diagnosis Method Based on Characteristics Frequency Ratio. *Sensors*, 20(5), 2020: 1519. https://doi.org/10.3390/s20051519

[7] H. Chen, S. Li, X. Lu, Q. Zhang, J. Zhu, J. Lu, Research on bearing fault diagnosis based on a multimodal method. *Mathematical Biosciences and Engineering*, 21(12), 2024: 7688-7706. https://doi.org/10.3934/mbe.2024338

[8] X. Zhang, Y. Liang, J. Zhou, Y. zang, A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM. *Measurement*, 69, 2015: 164-179. https://doi.org/10.1016/j.measurement.2015.03.017

[9] V. Vakharia, V.K. Gupta, P.K. Kankar, Efficient fault diagnosis of ball bearing using ReliefF and Random Forest classifier. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 39(8), 2017: 2969-2982. https://doi.org/10.1007/s40430-017-0717-9

[10] S. Huang, Z. Wang, J. Yang, T. Gong, Z. Shan, Y. Yang, Adaptive fast Walsh-Hadamard transform for magnetic flux leakage signal of broken wire damage extraction under noise background. *Nondestructive Testing and Evaluation*, 40(2), 2025: 564-584. https://doi.org/10.1080/10589759.2024.2325671

[11] Z. Zulfikar, S.A. Abbasi, A.R.M. Alamoud, Design of real time walsh transform for processing of multiple digital signals. *International Journal of Electrical and Computer Engineering*, 3(2), 2013: 197-206.

[12] T. Khanam, P.K. Dhar, S. Kowsar, J.-M. Kim, SVD-based image watermarking using the fast Walsh-Hadamard transform, key mapping, and coefficient ordering for ownership protection. *Symmetry*, 12(1), 2020: 52. https://doi.org/10.3390/sym12010052

[13] V. Dave, H. Thakker, V. Vakharia, Fault Identification of Ball Bearings using Fast Walsh Hadamard Transform, LASSO Feature Selection, and Random Forest Classifier. *FME Transactions*, 50(1), 2022: 202-209. https://doi.org/10.5937/fme2201202D

[14] M. Zhao, B. Tang, Q. Tan, Fault diagnosis of rolling element bearing based on S transform and gray level co-occurrence matrix. *Measurement Science and Technology*, 26(8), 2015: 085008. https://doi.org/10.1088/0957-0233/26/8/085008

[15] Y. Du, Y. Chen, G. Meng, J. Ding, Y. Xiao, Fault severity monitoring of rolling bearings based on texture feature extraction of sparse time–frequency images. *Applied Sciences*, 8(9), 2018: 1538. https://doi.org/10.3390/app8091538

[16] M. Robnik-Šikonja, I. Kononenko, Theoretical and empirical analysis of ReliefF and RReliefF. *Machine Learning*, 53, 2003: 23-69. https://doi.org/10.1023/A:1025667309714

[17] H.R. Kanan, K. Faez, An improved feature selection method based on ant colony optimization (ACO) evaluated on face recognition system. *Applied Mathematics and Computation*, 205(2), 2008: 716-725. https://doi.org/10.1016/j.amc.2008.05.115

[18] Z. Yin, J. Hou, Recent advances on SVM based fault diagnosis and process monitoring in complicated industrial processes. *Neurocomputing*, 174, 2016: 643-650. https://doi.org/10.1016/j.neucom.2015.09.081

[19] K. Shao, W. Fu, J. Tan, K. Wang, Coordinated approach fusing time-shift multiscale dispersion entropy and vibrational Harris hawks optimization-based SVM for fault diagnosis of rolling bearing. *Measurement*, 173, 2021: 108580. https://doi.org/10.1016/j.measurement.2020.108580

[20] J. Lu, W. Qian, S. Li, R. Cui, Enhanced K-nearest neighbor for intelligent fault diagnosis of rotating machinery. *Applied Sciences*, 11(3), 2021: 919. https://doi.org/10.3390/app11030919

[21] Y. Liu, Y. Cheng, Z. Zhang, J. Wu, Multi-information fusion fault diagnosis based on KNN and improved evidence theory. *Journal of Vibration Engineering & Technologies*, 10, 2022: 841-852. https://doi.org/10.1007/s42417-021-00413-8

[22] H.B. Ul Haq, R. Younis, M.S. Ali, Towards robust network security: Evaluating machine learning algorithms for intrusion detection. *Decision Making Advances*, 3(1), 2025: 12638. https://doi.org/10.31181/dma31202559

[23] D.-T. Hoang, H.-J. Kang, Rolling element bearing fault diagnosis using convolutional neural network and vibration image. *Cognitive Systems Research*, 53, 2019: 42-50. https://doi.org/10.1016/j.cogsys.2018.03.002

[24] S. Kavathekar, N. Upadhyay, P.K. Kankar, Fault classification of ball bearing by rotation forest technique. *Procedia Technology*, 23, 2016: 187-192. https://doi.org/10.1016/j.protcy.2016.03.016

[25] P. Konar, P. Chattopadhyay, Bearing fault detection of induction motor using wavelet and Support Vector Machines (SVMs). *Applied Soft Computing*, 11(6), 2011: 4203-4211. https://doi.org/10.1016/j.asoc.2011.03.014

[26] P.K. Kankar, S.C. Sharma, S.P. Harsha, Fault diagnosis of ball bearings using continuous wavelet transform. *Applied Soft Computing*, 11(2), 2011: 2300-2312. https://doi.org/10.1016/j.asoc.2010.08.011

[27] B. Li, P.-y. Liu, R.-x. Hu, S.-s. Mi, J.-p. Fu, Fuzzy lattice classifier and its application to bearing fault diagnosis. *Applied Soft Computing*, 12(6), 2012: 1708-1719. https://doi.org/10.1016/j.asoc.2012.01.020