An Improved Rolling Bearing Fault Diagnosis Model of Long Short-Term Memory Network Based on VMD Denoised Vibration Signals

Thomas Joseph and Sudeep U

Department of Mechanical Engineering, NSS College of Engineering, Palakkad, APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala, India. E-mail: tjk4411@gmail.com

Keerthi Krishnan K

Department of Electronics and Communication Engineering, NSS College of Engineering , Palakkad, APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala, India.

Sidra Khanam

Department of Mechanical Engineering, ZHCET, Aligarh Muslim University, Aligarh, India, 202002.

(Received 16 February 2024; accepted 25 July 2024)

Data driven fault diagnosis methods are increasingly being used for condition monitoring of rotating machinery in the era of Industry 4.0. The effectiveness of Variational Mode Decomposition (VMD) denoised vibration signals in improving the Long Short-Term Memory (LSTM) method for the intelligent fault diagnosis of a rolling bearing is reported. The raw and VMD denoised vibration signals of rolling bearings are provided as inputs to the LSTM network for classification of bearing condition. The efficacy of the methodology to extract the fault information is assessed through datasets obtained from experiment test rig and through the open-source dataset of Case Western Reserve University (CWRU). A comparative analysis is also carried out using both VMD based decomposition and denoising techniques along with four machine learning classifiers viz. Decision Tree, k-Nearest Neighbour (k-NN), Support Vector Machine (SVM) and Artificial Neural Network (ANN) by using the statistical features of the VMD modes. Among the different methods evaluated, VMD denoised signals when fed to LSTM result in the maximum classification accuracy of 99.14 % with experimental dataset. For the case of CWRU dataset, VMD denoised signals as input to the SVM resulted in maximum classification accuracy of 98.16 %.

1. INTRODUCTION

The timely detection of faults developed in rolling bearings is crucial in ensuring uninterrupted operation of rotating machinery without sudden unexpected failures. Automated rolling bearing fault detection methods using artificial intelligence have received considerable research interest recently. This has relevance in the evolution of autonomous manufacturing systems and the emergence of data driven decision making process for predictive maintenance.[1–4]

Studies show that machine learning5 and deep learning[6] methods play a vital role in automated fault identification of rolling bearings. In these methods, extraction of statistical fault features from acquired vibration signals is critical. Signal decomposition methods are employed by researchers for fault feature extraction.[7] Some of the signal decomposition methods including Empirical Mode Decomposition (EMD), Ensemble Empirical Mode Decomposition (EEMD) and Local Mean Decomposition (LMD) have been explored in the context of fault feature extraction of rolling bearing vibration signals.[8,9] The main problem with EMD method is the band mixing, which is eliminated by using EEMD.[8] In the case of the LMD method, the envelop spectrum of the product function is analyzed, there by detecting the inner and outer race faults in bearings in the presence of strong noise.[10] However, EMD, EEMD and LMD methods fail to identify the optimum number of modes without mode mixing which is addressed by VMD.[11] Many studies[12,13] have been proposed for improvements in VMD, especially for parameter tuning.

It is worth mentioning here that severe operating conditions, the presence of background noise, impulses and the vibration transmitted from nearby machine components often make the automated fault detection difficult.[14] For effective fault diagnosis, denoising the vibration signals before feature extraction is vital.[15] Several researchers have considered wavelet thresholding[16] where thresholding was performed on EMD modes for denoising. Automated denoising techniques using denoising autoencoder[17] and convolutional neural network (CNN)18 were also proposed. A two stage denoising method was implemented by choosing Intrinsic mode function (IMF) of VMD based on sample entropy and further using wavelet threshold denoising exhibited good results.[19] A denoising techniques was also implemented wherein after Single Value Decomposition (SVD), the VMD mode is removed considering the singular values and kurtosis.[20] However in both the reported studies, results are reported as envelope spectra of the denoised signal without fault classification.[21]

It is worth noting here that the kurtosis of the mode decom-

posed signal, which represents the randomness of the signal was used as an effective parameter for denoising.[16] The kurtosis value of an IMF which is rich in fault information of faulty bearings must be greater than the kurtosis of IMF of the signal from healthy bearings.

Many studies are reported in fault feature classification using machine learning methods including k-NN, ANN and SVM.[22–25] Few researchers have also proposed machine learning techniques including k-NN along with VMD[26] and Principal Component Analysis (PCA) along with SVM27 for fault feature extraction and classification.

The intelligent fault diagnosis methods use a pre-determined transformation and decomposition for transforming the time domain data to frequency domain. This requires larger computation time in feature extraction and classification. This is challenging in real time processing of the fault data which is essential in autonomous diagnostic systems. This difficulty is overcome by using the Long Short-Term Memory (LSTM) network.[9,28,29] A novel DWT-LSTM based fault diagnosis capable of discovering complex pattern from a large amount of data. This has resulted in improved fault detection accuracy.[30]

Even with the large number of studies on evaluation of performance of machine learning and deep learning classification of rolling bearing faults with various modal decomposition and denoising techniques, authors could not identify any specific study pertaining to the effectiveness of VMD in denoising the rolling bearing vibration signal and the efficacy of LSTM classifier for the intelligent classification of bearing fault condition. The objective of the present study is to explore the capability of the LSTM network in classifying rolling bearing fault conditions through VMD denoising of the time domain signals.

The raw time domain signals were acquired from test bearings categorized as healthy, with inner, outer race faults and with ball faults. The open data source of healthy and faulty rolling bearings available at the CWRU bearing data center was also employed in the present study for validation. The raw time domain signals were also directly fed to the LSTM network for comparison of performance of proposed methodology. The performance of the proposed method was further compared with two other VMD based methods for intelligent fault classification of components of rolling bearings using Decision Tree, k-NN, SVM and ANN. The classification accuracy, recall, precision and F1 scores were evaluated to compare the efficacy of different methods.

The remaining part of the article is structured into the following sections: Section 2 provides the theoretical background of VMD and LSTM methods. Section 3 describes the experimental details and the datasets used in the study. Section 4 provides the methodology adopted with a graphical description. Section 5 presents the results and the relevant discussions. Finally, Section 6 concludes the work by quantitatively highlighting the main findings with limitations and future scope.

2. THEORETICAL BACKGROUND

VMD is an optimization-based methodology for decomposing signals into frequency centered Intrinsic Mode Functions (IMFs) using the calculus of variation.[31–33] Each IMF is centered around a central frequency. An optimization methodology viz. Alternating Direction Method of Multipliers (ADMM) is used to identify the center frequency. The original formulation of the optimization problem is continuous in time domain, which is to minimize the sum of the bandwidths of all modes, subject to the condition that original signal is reconstructed by adding all the modes.

The mathematical expression is

$$\min_{u_k,\omega_k} \left\{ \sum_k \left\| \delta_t \left[\left((\sigma(t) + \frac{j}{\pi t}) * u_k(t) \right) e^{-j\omega_k t} \right] \right\|_2^2 \right. \quad \text{s.t.} \sum_k u_k = f; \quad (1)$$

where f is the original signal, λ - Lagrangian multipliers, $u_k - k^{th}$ IMF, ω_k - Centre frequency, σ- Dirac distribution, α - The balancing parameter of the data-fidelity constraint, t - Time-step of the dual ascent.

Augmented Lagrangian formulation for the above statement is as following,

$$L(u_k, w_k, \lambda) = \alpha \sum_k \left\| \delta_t \left[\left((\sigma(t) + \frac{j}{\pi t}) * u_k(t) \right) e^{-j\omega_k t} \right] \right\|_2^2 + \left\| f - \sum_k u_k \right\|_2^2 + \left\langle \lambda, f - \sum_k u_k \right\rangle; \quad (2)$$

In ADMM, one variable is solved at a time assuming that others are known and finally modes and central frequencies are updated using equation (3) until convergence condition is satisfied.

Update $\hat{u}_k$ for all $\omega \geq 0$

$$\hat{u}_k^{n+1} \leftarrow \frac{\hat{f} - \sum_{i<k} \hat{u}_i^{n+1} - \sum_{i>k} \hat{u}_i^n + \frac{\hat{\lambda}^n}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2}; \quad (3)$$

Update ω_k:

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega}. \quad (4)$$

2.1. LSTM Network

The basic architecture of the LSTM network classifier used for classification is shown in Fig. 1. Since the input data size is 20000, the LSTM network classifier's input layer comprises of 20000 neurons. These data points are then fed to the hidden layer which consists of 5000 LSTM cells. Each LSTM cell has an input gate, output gate, forget gate, and cell state. The cell state from previous LSTM cell and new acceleration information are sent to the adjacent LSTM cell. Since each acceleration data is presented in a succession of time series data, the LSTM network classifier operates on the premise that they are all interconnected. Finally, the output from hidden layer is given to the output layer for prediction. The "SoftMax" output layer provides the prediction as probabilities for the four conditions of rolling bearings. Further details of LSTM method used in this study are referred elsewhere.[34,35]

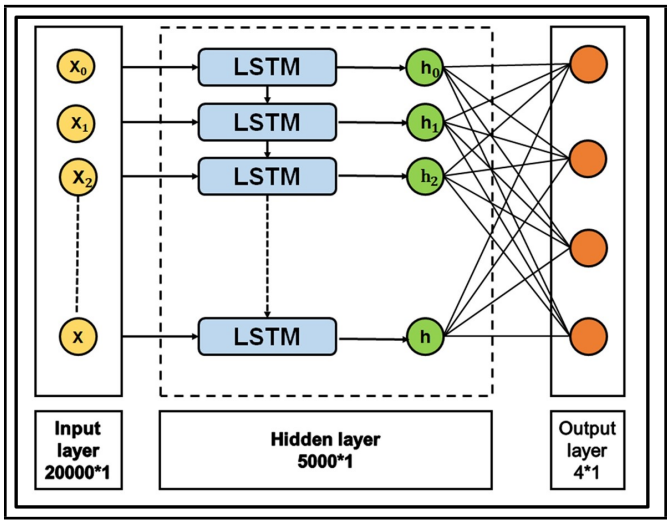

Figure 1. Basic architecture of LSTM network classifier.

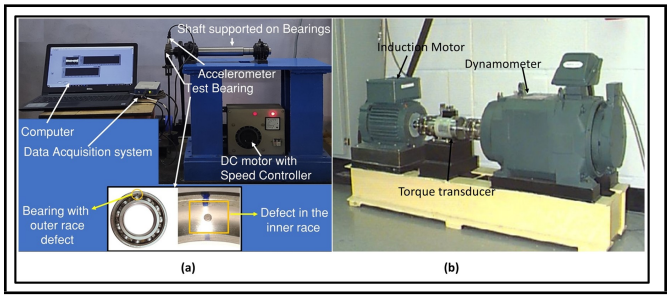

Figure 2. a) Experimental test rig for rolling bearing data acquisition (b) CWRU test rig.[36]

Table 1. Operating conditions and specifications of rolling bearing.

Parameters	Specifications
Operating Speed	1000 RPM
Shaft diameter	35 mm
DC Motor power	3 HP
Load range	100 N
Bearing type	SKF 6007
Number of balls and ball diameter	11 and 8.9 mm
Pitch diameter and contact angle	48.5 mm and 0°

3. DESCRIPTION OF EXPERIMENTAL AND CWRU DATASETS

The experimental vibration data were acquired from a setup consisting of a pulley driven shaft supported on two support bearings and a test bearing as shown in Fig. 2 (a). The operating conditions and specifications of the test bearing are indicated in Tab. 1. An ICP type of accelerometer was connected to the load zone of the test bearing, which in turn was connected to National Instruments (NI) 9250, which is a 2-channel, ± 5 V, vibration input module connected through a C-DAQ 9171 provided by NI to a computer for data logging and further processing. Experimental vibration data from the deep groove ball bearings classified as healthy, inner, outer races faults and ball faults have been obtained. Small sized circular faults measuring 1.0 mm diameter were seeded on the inner and outer bearing races and on the balls of different sets of bearings with the help of electric discharge machining. These faults simulated cracks and corrosion pitting at an early stage in the operational bearing's races and balls. The vibration data were acquired at a

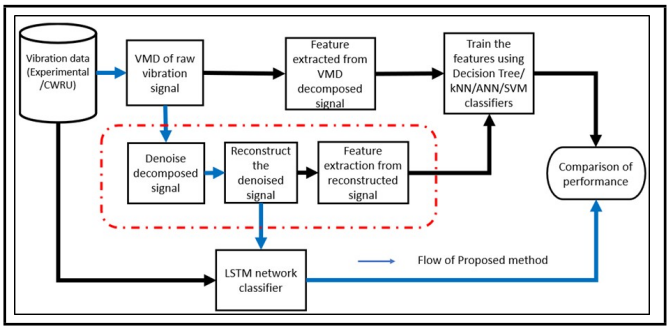

Figure 3. Flow diagram of the reported research work.

sampling rate of 20 kHz. The vibration signals were acquired for bearing conditions termed, healthy, inner, outer race fault and ball fault. For each condition of the bearing, 3000 sets of data of one second duration were acquired. Among the 3000 sets of data readings of each condition of the bearing, 80 % of the data were considered for training and 20 % were considered for testing of the LSTM and machine learning classification.

Open-source datasets available at the CWRU bearing data center[36] were used as the benchmark for validation of the proposed methodology. CWRU rolling bearing test rig is shown in Fig. 2 (b). The dataset provided by CWRU36 was acquired from a deep groove ball bearing with a fault of size 0.007, 014, and 0.021 inches on inner, outer race and balls of SKF bearings. The readings were taken under sampling frequencies 12000 Hz and 48000 Hz and a load of 0, 1 HP and 2 HP at various speeds. The accelerometer was placed at 6 o'clock, 3 o'clock and 12 o'clock position of outer race fault points.

The fault information from CWRU dataset at one HP load and 1772 rpm with a fault size of 0.007 inches under the sampling frequency of 12000 Hz were considered in this study. Here the load zone of the outer race fault was taken as 6 o'clock position.

The data were collected for a duration of 40 seconds for each condition of bearing. Thus, the study used each data of one second duration as one set of data. Hence among the 40 sets of data readings of each condition of bearing, 80 % data was used for training and 20 % for testing the LSTM and machine leaning algorithms.

4. FAULT CLASSIFICATION METHODOLOGY

Figure 3 shows the flow diagram of the proposed methodology for improving the LSTM classification of VMD denoised rolling bearing vibration signals. The raw time domain signals of the four bearing conditions along with the corresponding envelope spectra are shown in Fig. 4. Theoretically calculated fault frequencies[37] were Ball Pass Frequency Inner (BPFI) at 108.48 Hz, Ball Pass Frequency Outer (BPFO) at 74.84 Hz, and Ball Spin Frequency (BSF) at 43.88 Hz. The corresponding experimentally obtained frequencies at 109 Hz, 75 Hz, and 44 Hz respectively along with their harmonics are indicated in Fig. 4. The raw time domain vibration signals were first decomposed into various IMFs using VMD. While decomposing with VMD, the balancing parameter α plays a vital role in identifying the impacts of faults in bearings. If the value of α was small then VMD detected the impact of the faults.[38]

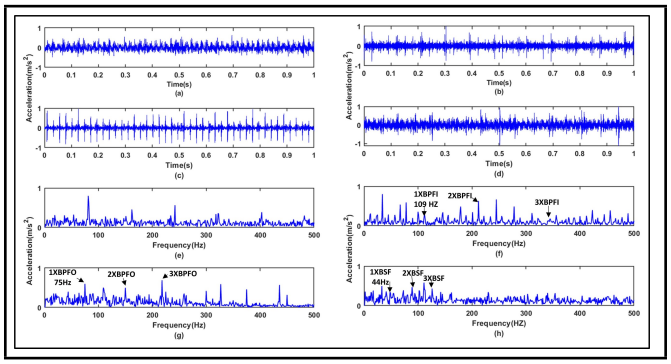

Figure 4. Raw time domain signals of (a-d) healthy, inner race fault, outer race fault and ball fault and (e-h) the corresponding envelope spectra.

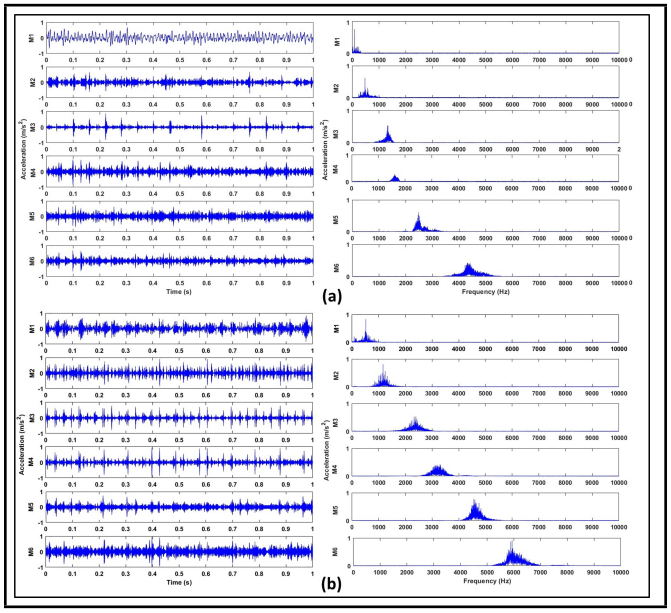

Figure 5. VMD decomposed signals of (a) healthy bearing (b) bearing with inner race fault.

Therefore, the balancing parameter was set as 2500 and the number of modes, K was set as 6.39 Representative VMD modes and the corresponding spectra of healthy and bearing with inner race fault are shown in Fig. 5. Thereafter, signals were VMD denoised based on the kurtosis parameter. For denoising, the IMF with kurtosis value less than three were discarded which carried the noise component of vibration signal. The signals were reconstructed from the remaining IMF.[40,41]

The reconstructed signals of bearing with healthy and the three fault conditions after denoising along with the corresponding envelope spectra are shown in Fig. 6. After denoising, the noise components are significantly reduced as observed from the envelope spectra.

4.1. Improved LSTM Method Of Classification

The denoised signals were fed as input to the LSTM classifier. The "Adam" optimizer was used to perform the optimization, and the "learning rate" was fixed at 0.001.

To have a comparison of performance, the raw vibration signals were also directly fed as input to the same LSTM network for fault classification. The procedure was repeated with the open dataset of CWRU for validation purpose.

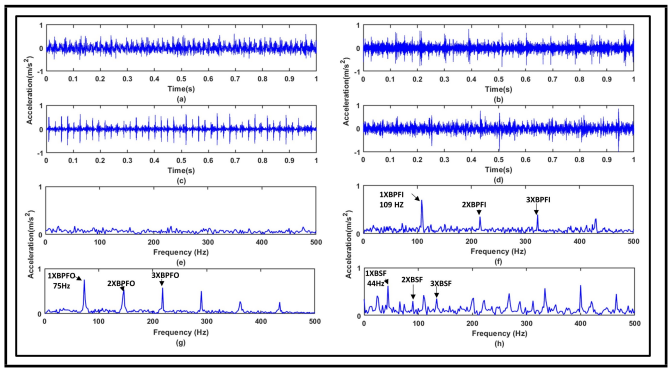

Figure 6. VMD denoised signals of (a) healthy bearing (b) bearing with inner race fault (c) bearing with outer race fault (d) bearing with ball fault and (e-h) the corresponding envelope spectra.

Table 2. Extracted statistical features.

Maxima	$\bar{Y}$ + Upper deviation
Minima	$\bar{Y}$ - Lower deviation
Mean	$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} Y_i$
Standard deviation	$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})}$
Skewness	$S_K = \frac{\sum_{i=1}^{N}(Y_i-\bar{Y})^3/N}{s^3}$
Kurtosis	$K = \frac{\sum_{i=1}^{N}(Y_i-\bar{Y})^3/N}{s^4}$
Entropy	$h(x) = \log P(x)$

Apart from the deep learning method i.e., LSTM, the effectiveness of VMD denoising was evaluated using four different machine learning classifiers viz. Decision Tree, k-NN, ANN and SVM after extracting the statistical features from both the VMD decomposed and VMD denoised vibration signals.

4.2. Classification By Decision Tree, K-NN, ANN And SVM

From each of the six VMD modes, seven statistical features viz. maxima, minima, mean, standard deviation, skewness, kurtosis and entropy were extracted.[42] The mathematical expressions of these features are provided in Tab. 2. Thus, from the vibration signals categorised as healthy, inner, outer race fault and ball fault, a total of seven features were extracted from each of the six modes which results in 42 features for each condition of bearing. Here N was the number of data points and $\bar{Y}$ was the mean and $P(x)$ was the probability of x.

The box plots indicating the variations of each statistical feature for the case of raw as well as the VMD denoised signals are shown in Fig. 7. Each box plot in subfigure indicates the variation of a statistical feature which were normalised and indicated using a colour scheme. Figure 7 (a-d) indicates the variation of raw signals for the four conditions of bearing, while Fig. 7 (e-h) indicates the variations of VMD denoised signals for the four conditions of bearing using the experimental dataset. Box plots were generated considering all the samples in the dataset. The mean value of each of the statistical parameter remained unchanged after denoising, however the maxima and minima values of each parameter is observed to have variations in magnitude. Similarly, Fig. 7 (i-l) and Fig. 7 (m-p) shows the variation of statistical features for the cases of four classes of bearing conditions of the CWRU dataset. For visualizing the effect of denoising on the distribution of features, t-SNE plots are shown in Fig. 8 for the case

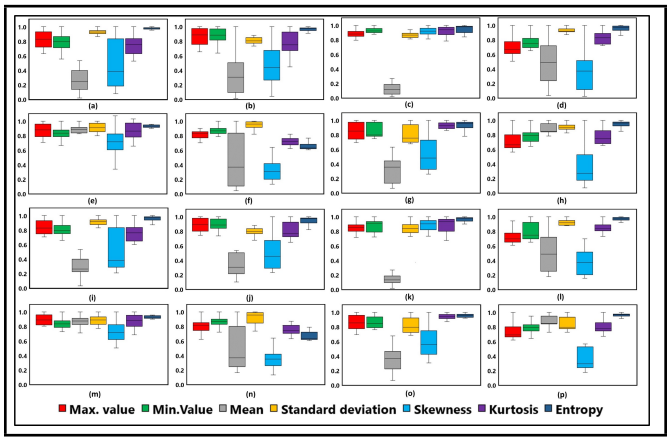

Figure 7. Box plots of statistical features of raw signals of (a-d) and (i-l) healthy, inner, outer race faults and ball fault for experimental dataset and of CWRU dataset respectively (e-h) and (m-p) corresponding conditions of the VMD denoised signal.

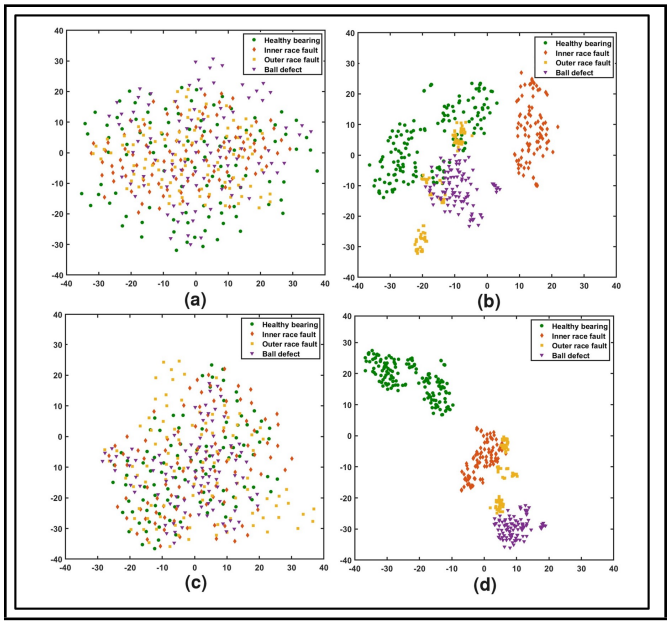

Figure 8. t-SNE plot of (a-b) raw vibration signal and VMD denoised signals of experimental dataset (c-d) raw vibration signal and VMD denoised signals of CWRU datasets.

of experimental and CWRU datasets. Here Fig. 8 (a) and (b) represent the features of experimental dataset before and after denoising and Fig. 8 (c) and (d) represent the corresponding features of CWRU dataset before and after denoising. Each colour represented the datapoints corresponding to each condition of bearing projected to the two dimensions from the seven dimensions of the statistical features. Figures 8 (a) and (b) provides the t-SNE plots for the raw and denoised data from the experiments. Here the separation of four bearing conditions is visible after denoising. For the case of CWRU dataset, the separation of the four bearing conditions is more evident as indicated in Figs. 8(c) and (d).

5. RESULTS AND DISCUSSION

5.1. Results Of LSTM Classifier

Accuracy and loss plots of LSTM network classifier using the raw and VMD denoised signals are shown in Fig. 9. for the

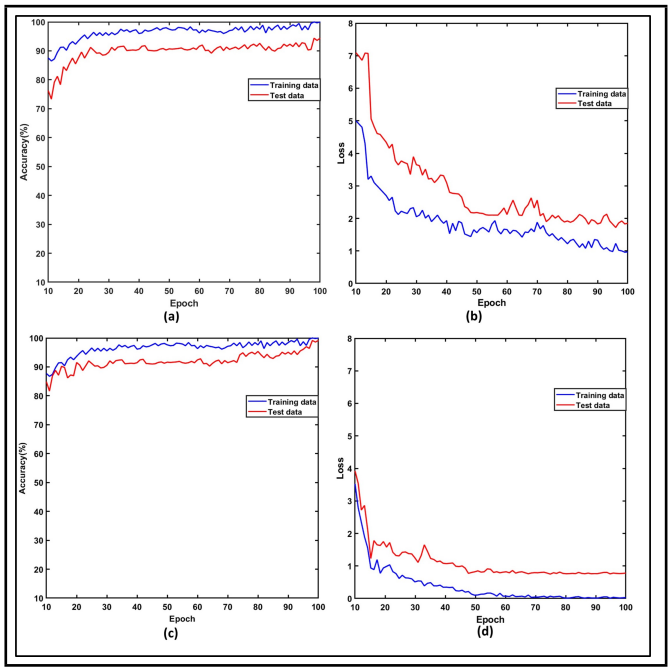

Figure 9. LSTM network classifier accuracy plot of (a) raw signal (b)VMD denoised signal and loss plot of (c) raw signal (d) VMD denoised signal of experimental dataset.

Table 3. Classification accuracy using the LSTM method.

Dataset/Method		Classification accuracy (%)	Recall	Precision	F1 Score
Experimental dataset	Raw signal	94.37	0.9433	0.9982	0.9700
	VMD-denoised signal	99.14	0.9916	0.9983	0.9949
CWRU dataset	Raw signal	82.00	0.9899	0.9879	0.9889
	VMD-denoised signal	87.30	0.9905	0.9887	0.9895

case of experimental dataset. The results in terms of classification accuracy, recall, precision and F1 scores for the case of raw and VMD denoised signals are provided in Tab. 3. The raw time domain signals when fed to the LSTM network resulted in a maximum classification accuracy of 94.37 % whereas the classification accuracy of 99.14 % was obtained in the case of VMD denoised signals using the experimental dataset. For the case of CWRU dataset an accuracy of 82 % and 87.30 % were obtained respectively for the raw and VMD denoised signals. The results reported in Tab. 3 indicate that the accuracy, recall, precision and F1 scores of VMD-denoised signals are higher in comparison with the raw signal for the case of experimental and CWRU datasets. This indicates the effectiveness of VMD denoising along with the LSTM classifier.

5.2. Results Of Decision Tree, K-NN, ANN And SVM Classifiers

To evaluate the efficacy of denoising technique using the different machine learning methods, statistical features extracted from both the VMD decomposed and VMD denoised signals were classified using the Decision Tree, k-NN, SVM and ANN classifiers. The confusion matrix obtained by classifying the VMD decomposed and VMD denoised signals are shown in Fig. 10 and 11 respectively. The methodology of classification by using the features of VMD decomposed signals and VMD denoised signals were further validated using the CWRU

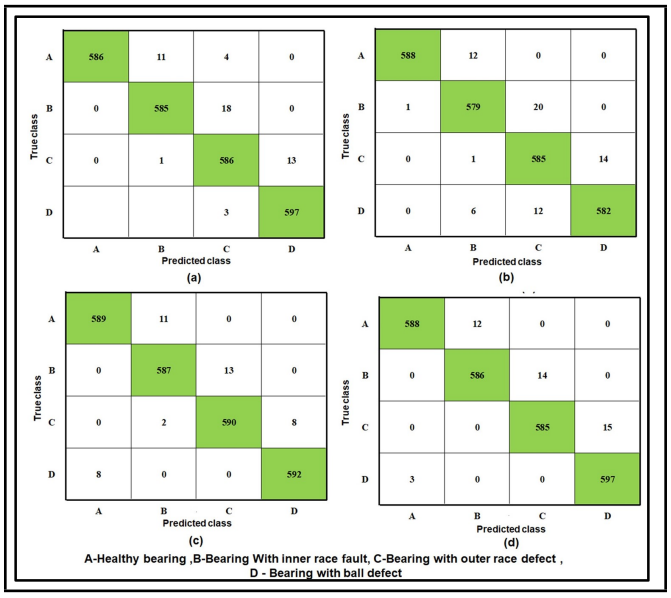

Figure 10. Confusion matrix of (a) Decision Tree (b) k-NN (c) SVM and (d)ANN classifier of VMD decomposed signal of experimental dataset.

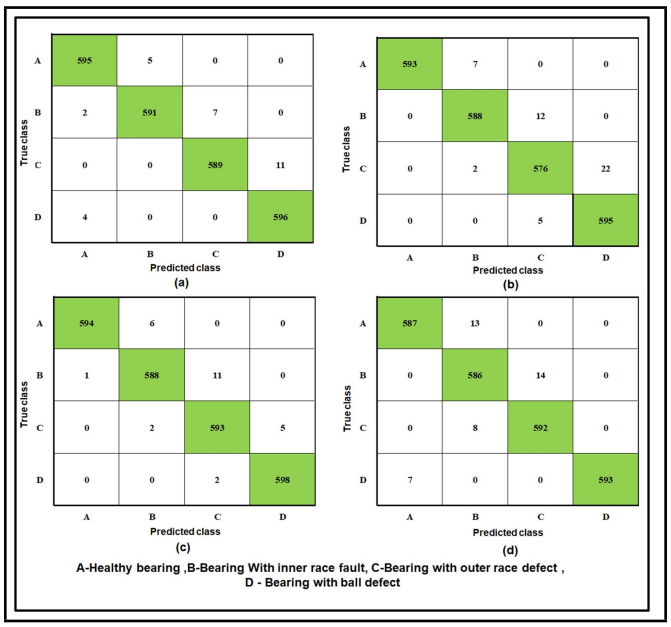

Figure 11. Confusion matrix of (a) Decision Tree (b) k-NN (c) SVM and (d)ANN classifier of VMD denoised vibration signal of experimental dataset.

dataset. Table 4 shows the condition wise classification accuracy with four different classifiers and two datasets. The inner race fault classification accuracy was less compared to the outer race and ball fault condition, while the ball fault and the outer race faults were classified with better accuracy while using both the datasets. Results using the experimental dataset show that the condition wise accuracy of Decision Tree classifier was highest in classifying the ball fault and the ANN classifier showed the maximum class wise accuracy in classifying the outer race fault. The tabulated results show that the class wise classification accuracy using VMD denoised signals were the highest for the case of both the datasets used herein. The classification accuracy using VMD denoised signals with SVM showed a class wise accuracy value of 98.83 % while classifying the outer race fault and a value of 99.6 % while classifying the ball faults using the experimental dataset. This method

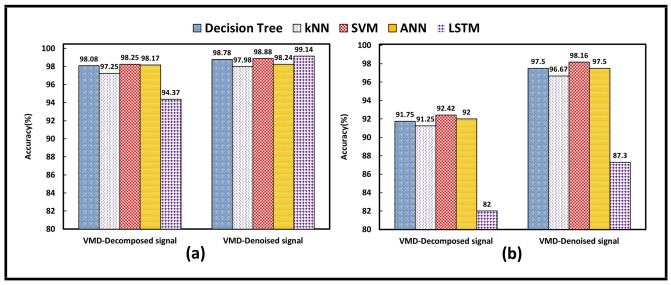

Figure 12. Overall classification accuracy for VMD decomposed and VMD denoised signals (a) Experimental dataset (b) CWRU dataset.

showed the highest accuracy in classifying the ball faults.

Table 5 provides evaluation parameters of Machine Learning (ML) such as classification accuracy, recall, precision and F1 scores. The results in Tab. 5 show that the SVM classifier provided maximum classification accuracy of 98.25 % for VMD decomposed signal and an accuracy of 98.88 % for the VMD denoised signal using the experimental dataset while 92.42 % and 98.16 % accuracies using the CWRU dataset for respective conditions.

The overall accuracy of VMD decomposed and VMD denoised signals of experimental and CWRU datasets for the case of four ML classifiers and LSTM classifier are shown in Fig. 12. It is observed that the SVM classifier outperformed all other ML classifiers and LSTM Classifier outperformed for experimental signals after VMD denoising. Significant improvements in classification accuracy were observed for all the classifiers by using the VMD denoised signal. The recall value of VMD-denoised LSTM as provided in Tab. 3 showed maximum value in comparison with Decision Tree, k-NN, SVM and ANN classifiers which indicate that VMD-denoised LSTM classifier outperformed other classifiers in classifying the condition of rolling bearing. The precision value of VMD denoised LSTM classifier reached the maximum at 0.9983. The F1 score of the VMD denoised LSTM classifier showed maximum among all the classifiers for the raw and VMD decomposed vibration signals.

Since among the ML classifiers, SVM provided best classification accuracy for both the datasets, a comparison of classification accuracy of SVM and LSTM classifiers are provided in Tab. 6. It is worth noting here that classification accuracy remained consistently higher for the case of VMD denoised signal for both the SVM and LSTM classifier with the datasets used in this study. For the case of experimental dataset by using LSTM with VMD denoised signals as input provided higher accuracy of 99.14 % compared to SVM. However, for CWRU dataset SVM, provided the better accuracy of 98.16 % for denoised signals. It is observed that LSTM classifier resulted in better classification accuracy with VMD denoised signals when a greater number of samples were available from experimental dataset. In contrast, when the number of samples were limited, SVM resulted in better classification accuracy with VMD denoised signals. It is evident that deep learning classifier performed better with larger data samples and machine learning classifier performed better with limited number of sampled vibration data. To compare the present work with the state-of-the-art results, Tab. 7 provides the details of features used, classifier and the classification accuracy reported in

Table 4. Comparison of class wise accuracy using experimental and CWRU datasets.

Classifier/condition of bearing		Decision Tree				k-NN				ANN				SVM			
		Healthy	Inner race fault	Outer race fault	Ball fault	Healthy	Inner race fault	Outer race fault	Ball fault	Healthy	Inner race fault	Outer race fault	Ball fault	Healthy	Inner race fault	Outer race fault	Ball fault
Experimental dataset	VMD decomposed	97.66	97.50	97.66	99.50	98.00	96.50	97.50	97.00	98.16	97.83	98.33	98.66	98.00	97.66	97.50	99.50
	VMD denoised	99.17	98.50	98.10	99.33	98.83	98.00	96.00	99.10	99.00	98.00	98.83	99.60	97.83	97.67	98.66	98.83
CWRU dataset	VMD decomposed	92.33	90.67	91.67	92.33	91.33	89.67	93.33	90.67	92.32	94.52	90.62	92.22	93.67	91.33	92.33	90.67
	VMD denoised	100	96.67	93.33	100	100	93.33	93.33	100	100	97.20	99.12	96.32	100	96.67	96.67	96.67

Table 5. Classification accuracy in percentage of all classification methods.

Dataset	Method	Classifier	Classification accuracy (%)	Recall	Precision	F1 Score
Experimental dataset	VMD decomposed signal	Decision Tree	98.08	0.9792	0.9794	0.9792
		k-NN	97.25	0.9725	0.9728	0.9726
		SVM	98.25	0.9825	0.9825	0.9825
		ANN	98.17	0.9817	0.9817	0.9817
	VMD denoised signal	Decision Tree	98.79	0.9879	0.9879	0.9879
		k-NN	98.00	0.9800	0.9802	0.9800
		SVM	**98.88**	**0.9888**	**0.9888**	**0.9888**
		ANN	98.25	0.9825	0.9826	0.9825
CWRU dataset	VMD decomposed signal	Decision Tree	91.75	0.9250	0.9307	0.9254
		k-NN	91.25	0.9243	0.9288	0.9239
		SVM	92.42	0.9375	0.9452	0.9385
		ANN	92.00	0.9250	0.9313	0.9253
	VMD denoised signal	Decision Tree	97.50	0.9750	0.9733	0.9753
		k-NN	96.67	0.9625	0.9631	0.9625
		SVM	98.16	0.9875	0.9881	0.9875
		ANN	97.50	0.9750	0.9756	0.9750

Table 6. Comparison of classification accuracy.

Method	Accuracy of classification using experimental dataset (%)	Accuracy of classification using CWRU dataset (%)
Raw Signal +LSTM	94.37	82.00
VMD denoised +LSTM	**99.14**	87.30
VMD decomposed +SVM	98.25	92.42
VMD denoised +SVM	98.88	**98.16**

the recent literature. In comparison with the reported results, VMD denoising of rolling bearing vibration signal has resulted in overall improvement by using LSTM and SVM classifiers for experiment dataset and CWRU dataset respectively.

6. CONCLUSIONS

This research work addresses the need for improving the performance of automatic fault diagnosis rolling bearing faults in noisy environment using data driven methods. An improved methodology of LSTM classification by employing VMD denoising is reported. Experimentally obtained vibration signals from healthy and bearings with seeded faults on races and balls are denoised by VMD based on the kurtosis parameter. Envelope spectra of vibration signals are obtained before and after

Table 7. Summary of recent results reported in the literature.

Reference	Features used	Classifier	Maximum classification accuracy
43	Multi-domain feature set containing features from time domain, frequency domain,and entropy energy	SVM	100 %
44	Amplitude, BPFO, Amplitude, BPFI	Fine KNN	100 %
45	Discrete Wavelet Transform, Expectation Selection Maximization method of feature selection	GMM	80 %
46	Spectrogram as 224-pixel square RGB images	ResNet trained and tested using different set of data	99.5 %

denoising which indicated the extent of denoising. The denoised signals are fed to the LSTM classifier for fault classification. A classification accuracy of 99.14 % is reported. Raw time domain signals are also fed directly to the LSTM for comparison purposes which resulted in classification accuracy of 94.37 %. In addition to the LSTM classifier, four machine learning classifiers viz. Decision Tree, k-NN, SVM and ANN are also employed to evaluate the efficacy of VMD denoising. Recall, precision and F1 scores are also calculated and compared for LSTM and all machine learning classifiers employed. All the machine learning classifiers achieved higher fault classification accuracy for VMD denoised signals compared to the VMD decomposed signals. SVM classifier provided maximum classification accuracy of 98.88 % among machine learning classifiers. The methodology is further evaluated with the benchmark open dataset of CWRU. In contrast to the results obtained using the experimental dataset, the best classification accuracy of 98.16 % is observed by SVM classifier using the VMD denoised signal.

Overall, it is found that VMD denoising improved classification accuracy in comparison with raw signals and VMD decomposed signals. LSTM classifier provided best classification accuracy using experimental dataset whereas SVM classifier provided best accuracy using dataset of CWRU. The proposed methodology still has some limitations since it has considered only kurtosis parameter for VMD denoising. This limitation can be overcome by employing other correlation parameters for denoising the signals, which is the scope for future work.

REFERENCES

[1] Hakim, M., Omran, A. A. B., Ahmed, A. N., Al-Waily, M. and Abdellatif, A. A systematic review of rolling bearing fault diagnoses based on deep learning and transfer learning: Taxonomy overview application open challenges weaknesses and recommendations, *Ain Shams Engineering Journal.*, **14**, 101945, (2023). https://doi.org/10.1016/j.asej.2022.101945

[2] Fan, C. Peng, Y. Shen, Y. Guo, Y. Zhao, S. Zhou, J. and Li, S. Variable scale multilayer perceptron for helicopter transmission system vibration data abnormity beyond efficient recovery, *Eng Appl Artif Intell.*, **133**, 108184, (2024). https://doi.org/10.1016/j.engappai.2024.108184

[3] Li, S., Peng, Y. and Bin, G. Prediction of wind turbine blades icing based on CJBM with imbalanced data, *IEEE Sens J.*, **23**, 19726–19736, (2023). https://doi.org/10.1109/JSEN.2023.3296086

[4] Li, S. Peng, Y. and Bin, G. Rolling bearing fault diagnosis under data imbalance and variable speed based on adaptive clustering weighted oversampling, *Reliab Eng Syst Saf.*, **244**, 109938, (2024). https://doi.org/10.1016/j.ress.2024.109938

[5] Cui, M., Wang, Y., Lin, X. and Zhong, M. Fault diagnosis of rolling bearings based on an improved stack autoencoder and support vector machine, *IEEE Sens J.*, **21**, 4927–4937, (2021). https://doi.org/10.1109/JSEN.2020.3030910

[6] Mushtaq, S., Islam, M. M. M. and Sohaib, M. Deep learning aided data-driven fault diagnosis of rotatory machine: A comprehensive review, *Energies (Basel).*, **14**, 5150, (2021). https://doi.org/10.3390/en14165150

[7] Zhang, C., Jin, R. and Li, C. Fault diagnosis of synchronous generator based on adaptive chirp mode decomposition permutation entropy and deep learning, *Int. J. Acoust. Vib.*, **28**, 371–380, (2023). https://doi.org/10.20855/ijav.2023.28.41977

[8] Keshtan, M. N. and Nouri Khajavi, M. Bearings fault diagnosis using vibrational signal analysis by EMD method, *Res. Nondestruct. Evaluation*, **27**, 155–174, (2016). https://doi.org/10.1080/09349847.2015.1103921

[9] Zou, P., Hou, B., Jiang, L. and Zhang, Z. Bearing fault diagnosis method based on EEMD and LSTM, *Int. J. Comput. Commun. Control.*, **15**, (2020). https://doi.org/10.15837/ijccc.2020.1.3780

[10] Li, Y., Si, S., Liu, Z. and Liang, X. Review of local mean decomposition and its application in fault diagnosis of rotating machinery, *J. Syst. Eng. Electron.*, **30**, 799–814, (2019). https://doi.org/10.21629/JSEE.2019.04.17

[11] Zhang, M., Jiang, Z. and Feng, K. Research on variational mode decomposition in rolling bearings fault diagnosis of the multistage centrifugal pump, *Mech Syst Signal Process.*, **93**, 460–493, (2017). https://doi.org/10.1016/j.ymssp.2017.02.013

[12] Li, H., Wu, x., Liu, T., Li, S., Zhang, B., Zhou, G. and Huang, T. Composite fault diagnosis for rolling bearing based on parameter-optimized VMD, *Measurement.*, **201**, 111637, (2022). https://doi.org/10.1016/j.measurement.2022.111637

[13] Zhang, X., Miao, Q., Zhang, H. and Wang, L. A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery, *Mech Syst Signal Process.*, **108**, 58–72, (2018). https://doi.org/10.1016/j.ymssp.2017.11.029

[14] Liang, P., Wang, W., Yuan, X., Liu, s., Zhang, L. and Cheng, Y. Intelligent fault diagnosis of rolling bearing based on wavelet transform and improved ResNet under noisy labels and environment, *Eng Appl Artif Intell.*, **115**, (2022). https://doi.org/10.1016/j.engappai.2022.105269

[15] Chen, W., Li, J., Wang, Q. and Han, K. Fault feature extraction and diagnosis of rolling bearings based on wavelet thresholding denoising with CEEMDAN energy entropy and PSO-LSSVM, *Measurement.*, **172**, 108901, (2021). https://doi.org/10.1016/j.measurement.2020.108901

[16] Yin, C., Wang, Y., Ma, G., Wang, Y., Sun, Y. and Yan, H. Weak fault feature extraction of rolling bearings based on improved ensemble noise-reconstructed EMD and adaptive threshold denoising, *Mech Syst Signal Process.*, **171**, 108834. (2022). https://doi.org/10.1016/j.ymssp.2022.108834

[17] Che, C., Wang, H., Ni, X. and Fu, Q. Intelligent fault diagnosis method of rolling bearing based on stacked denoising autoencoder and convolutional neural network, *Ind. Lubr. Tribol.*, **72**, 947–953, (2020). https://doi.org/10.1108/ILT-11-2019-0496

[18] Han, H., Wang, H., Liu, Z. and Wang, J. Intelligent vibration signal denoising method based on non-local fully convolutional neural network for rolling bearings, *ISA Trans.*, **122**, 13–23, (2022). https://doi.org/10.1016/j.isatra.2021.04.022

[19] Ma, J., Li, H., Chen, Y., Wang, J. and Zou, Z. Application of VMD and dynamic wavelet noise reduction techniques in rolling bearing fault diagnosis, *J Phys Conf Ser.*, **2528**, 012048, (2023). https://doi.org/10.1088/1742-6596/2528/1/012048

[20] Wang, Q., Wang, L., Yu, H., Wang, D. and Nandi, A. K. Utilizing SVD and VMD for denoising non-stationary signals of roller bearings, *Sensors*, **22**, 195, (2021). https://doi.org/10.3390/s22010195

[21] Li, J., Wang, H., Wang, X. and Zhang, Y. Rolling bearing fault diagnosis based on improved adaptive parameterless empirical wavelet transform and sparse denoising, *Measurement*, **152**, 107392, (2020). https://doi.org/10.1016/j.measurement.2019.107392

[22] Moosavian, A., Ahmadi, H., Tabatabaeefar, A. and Khazaee, M. Comparison of two classifiers; K-nearest neighbor and artificial neural network, for fault diagnosis on a main

engine journal-bearing, *Shock. Vib.*, **20**, 263–272, (2013). https://doi.org/10.3233/SAV-2012-00742

[23] Soualhi, A., Medjaher, K. and Zerhouni, N. Bearing health monitoring based on Hilbert-Huang transform, support vector machine and regression, *IEEE Trans Instrum Meas.*, **64**, 52–62, (2015). https://doi.org/10.1109/TIM.2014.2330494

[24] Jamil, M. A., Khan, M. A. A. and Khanam, S. Feature-based performance of SVM and KNN classifiers for diagnosis of rolling element bearing faults, *Vibroengineering Procedia*, **39**, 36–42, (2021). https://doi.org/10.21595/vp.2021.22307

[25] Jamil, M. A. and Khanam, S. Influence of one-Way ANOVA and Kruskal–Wallis based feature ranking on the performance of ML classifiers for bearing fault diagnosis, *J. Vib. Eng. Technol.*, **12**, 3101-3132, (2024). https://doi.org/10.1007/s42417-023-01036-x

[26] Lu, Q., Shen, X,.Wang, X. and Li, M. Fault Diagnosis of rolling bearing based on improved VMD and KNN, *Math Probl. Eng* ., **2021**, 1-11, (2021). https://doi.org/10.1155/2021/2530315

[27] Deng, L., Zhang, A. and Zhao, R. Intelligent identification of incipient rolling bearing faults based on VMD and PCA-SVM, *Adv. Mech. Eng.*, **14**(1), (2022) https://doi.org/10.1177/16878140211072990

[28] Yang, M., Liu, W., Zhang, W., Wang, M. and Fang, X. Bearing vibration signal fault diagnosis based on LSTM-Cascade CatBoost, *J. Internet Technol.*, **23**, 1155–1161, (2022). https://doi.org/10.53106/160792642022092305024

[29] Kumar, A., Parey, A. and Kankar, P. K. A new hybrid LSTM-GRU model for fault diagnosis of polymer gears using vibration signals, *J. Vib. Eng. Technol.*, **12**(2), (2023). https://doi.org/10.1007/s42417-023-01010-7

[30] Gu, K., Zhang, Y., Liu, X., Li, H. and Ren, M. DWT-LSTM-based fault diagnosis of rolling bearings with multi-sensors, *Electronics (Switzerland).*, **10**, (2021). https://doi.org/10.3390/electronics10172076

[31] Krishnan K., K. and Soman, K. P. CNN based classification of motor imaginary using variational mode decomposed EEG-spectrum image, *Biomed Eng Lett.*, **11**, 235–247 (2021) https://doi.org/10.1007/s13534-021-00190-z

[32] Dragomiretskiy, K. and Zosso, D. Variational mode decomposition, *IEEE Trans. Signal Process.*, **62**, 531–544, (2014). https://doi.org/10.1109/TSP.2013.2288675

[33] Krishnan, K. K. and Soman, K. P. Comparison of variational mode decomposition and empirical wavelet transform methods on EEG signals for motor imaginary applications, I*nt J Biomed Eng Technol.*, **38**, 267, (2022). https://doi.org/10.1504/IJBET.2022.121740

[34] Liu, Z. H., Meng, X. D., Wei, H. L., Lu, B. L., Wang, Z. H. and Chen, L. A Regularized LSTM method for predicting remaining useful life of rolling bearings, *Int. J. Autom. Comput.*, **18**, 581–593, (2021). https://doi.org/10.1007/s11633-020-1276-6

[35] Li, J., Wang, Z., Liu, X. and Feng, Z. Remaining useful life prediction of rolling bearings using GRU-DeepAR with adaptive failure threshold, *Sensors*, **23**, (2023) https://doi.org/10.3390/s23031144

[36] Case Western Reserve University Bearing Data Center Website, `csegroups.case.edu/bearingdatacenter/home`, (Accessed November 15, 2023).

[37] Harris, T. A. and Kotzalas, *M. N. Rolling Bearing Analysis - 2 Volume Set*, CRC Press, (2006). https://doi.org/10.1201/9781482275148

[38] Wang, Y. and Markert, R. Filter bank property of variational mode decomposition and its applications, *Signal Processing.*, **120**, 509–521, (2016). https://doi.org/10.1016/j.sigpro.2015.09.041

[39] Wang, Y., Markert, R., Xiang, J. and Zheng, W. Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system, *Mech Syst Signal Process.*, **60**, 243–251, (2015). https://doi.org/10.1016/j.ymssp.2015.02.020

[40] Zhang, X., Luan, Z. and Liu, X. Fault diagnosis of rolling bearing based on kurtosis criterion VMD and modulo square threshold, *The J. Eng.*, **2019**, 8685–8690, (2019). https://doi.org/10.1049/joe.2018.9084

[41] Jie, D., Zheng, G., Zhang, Y., Ding, X. and Wang, L. Spectral kurtosis based on evolutionary digital filter in the application of rolling element bearing fault diagnosis, *Int. J. Hydromechatronics.*, **4**, 27, (2021). https://doi.org/10.1504/IJHM.2021.114173

[42] Sugumaran, V. and Ramachandran, K. I. Effect of number of features on classification of roller bearing faults using SVM and PSVM, *Expert Syst Appl.*, **38**, 4088–4096, (2011). https://doi.org/10.1016/j.eswa.2010.09.072

[43] Zhang, X., Zhao, B. and Lin, Y. Machine learning based bearing fault diagnosis using the Case Western Reserve University Data A Review, *IEEE Access*, **9**, 155598–155608, (2021). https://doi.org/10.1109/ACCESS.2021.3128669

[44] Alonso-Gonzalez, M., Diaz, V. G., Lopez Perez, B., Cristina Pelayo G-Bustelo, B. and Anzola, J. P. Bearing fault diagnosis with envelope analysis and machine learning approaches using CWRU dataset, *IEEE Access*, **11**, 57796–57805, (2023). https://doi.org/10.1109/ACCESS.2023.3283466

[45] Chaleshtori, A. E. and Aghaie, A. A novel bearing fault diagnosis approach using the Gaussian mixture model and the weighted principal component analysis, *Reliab Eng Syst Saf.*, **242**, 109720. (2024). https://doi.org/10.1016/j.ress.2023.109720

[46] Hendriks, J., Dumond, P. and Knox, D. A. Towards better benchmarking using the CWRU bearing fault dataset, *Mech Syst Signal Process*., **169**, (2022). https://doi.org/10.1016/j.ymssp.2021.108732