Journal of Systems Engineering and Electronics
Vol. 33, No. 1, February 2022, pp.233 – 246

Deep neural network based classification of rolling element bearings and health degradation through comprehensive vibration signal analysis

KULEVOME Delanyo Kwame Bensah[1,2], WANG Hong[1,2,*], and WANG Xuegang[1]

1. School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China; 2. Yangtze Delta Region Institute (Huzhou), University of Electronic Science and Technology of China, Huzhou 313001, China

Abstract: Rolling element bearings are machine components used to allow circular movement and hence deliver forces between components of machines used in diverse areas of industry. The likelihood of failure has the propensity of increasing under prolonged operation and varying working conditions. Hence, the accurate fault severity categorization of bearings is vital in diagnosing faults that arise in rotating machinery. The variability and complexity of the recorded vibration signals pose a great hurdle to distinguishing unique characteristic fault features. In this paper, the efficacy and the leverage of a pre-trained convolutional neural network (CNN) is harnessed in the implementation of a robust fault classification model. In the absence of sufficient data, this method has a high-performance rate. Initially, a modified VGG16 architecture is used to extract discriminating features from new samples and serves as input to a classifier. The raw vibration data are strategically segmented and transformed into two representations which are trained separately and jointly. The proposed approach is carried out on bearing vibration data and shows high-performance results. In addition to successfully implementing a robust fault classification model, a prognostic framework is developed by constructing a health indicator (HI) under varying operating conditions for a given fault condition.

Keywords: bearing failure, deep neural network, fault classification, health indicator, prognostics and health management.

DOI: 10.23919/JSEE.2022.000023

1. Introduction

During the operating life of a machine, its parts undergo various degrees of stress under varying conditions. In rotating machinery, bearings are used in the rotor and shaft of the rotor as a form of support and energy conversion. The likelihood of fault emanating from continuous operation increases with time. Machine vibration or noise levels, whether excessive or not, are affected by bearings in certain aspects [1]. It has been estimated that, about half of the faults that occur in rotating machinery originate from a rolling bearing failure [2] which poses a great hindrance to the safety and reliability of such machinery. Therefore, to guard against a sudden breakdown in such an equipment leading to huge losses, it is necessary to put in place a mechanism to effectively detect faults when they occur and monitor their degeneracy.

Diagnosis and monitoring of progressive bearing deterioration in rotating machinery by means of vibration measurements have been in use for some time and have become more economical and reliable in recent years [3–5]. Understanding the vibration signals is fundamental to effective fault detection and diagnostics. Therefore, three approaches mainly used over the years have been the analysis of vibration signals in time, frequency and time-frequency domains. In time domain representation of vibration signals, higher amplitudes beyond a normal level indicates the inception of a fault [5]. Traditional methods relying on the extraction of statistical features have been extensively studied over the past 20 years [6–8]. Techniques such as time synchronous averaging [9], autoregressive modeling [10], and blind source separation [11] have been explored in research with appreciable accuracies and associated challenges. However, it has been realized from the review of recent studies [12] that, time domain analysis has the ability to only indicate the presence of bearing faults but not the location of the fault. Consequently, researchers developed a couple of frequency domain techniques in analyzing vibration signals in bearings. Prominent among these techniques are the use of envelope analysis [13] and frequency domain feature engineering [14]. In similar studies, Tsao et al. [15]

Manuscript received November 10, 2020.
*Corresponding author.
This work was supported by the National Natural Science Foundation of China (42027805), National Aeronautical Fund (ASFC-2017 2080 005) and National Key R&D Program of China (2017YFC03 07100).

proposed an empirical mode decomposition method for selecting appropriate intrinsic mode functions (IMFs) for envelope analysis. However, due to the frequent variation in the operating speed of rotating machines, their vibration signals become nonstationary [16]. This gave rise to extensive research in processing vibration signals in the time-frequency domain for fault diagnosis in such machinery. Considerable among these techniques include the short-time Fourier transform [17], wavelet analysis [18], empirical mode decomposition [19], and Hilbert-Huang transform [20]. These approaches have achieved some level of success. However, bereft of prior knowledge, it becomes increasingly difficult to choose the features to extract. This has led to fault diagnosis of bearings using intelligent techniques as it provides intuitive diagnosis outcomes in the absence of extensive prior knowledge [21].

The advancement of machine learning (ML) and deep learning (DL) has propelled research in their application in machine health diagnosis by automatic extraction of features. The performance of these methods is highly dependent on the accessibility of immense data which must comprise healthy and all possible fault conditions. This is highly unlikely as machines are operated in their normal conditions, hence, there is large data represented by normal operating condition with little or none for faulty ones. For this reason, varied faults are simulated in the laboratory on bearings to collect ample labeled data analogous to real time operating conditions. This makes it possible to formulate an intelligent diagnosis framework for the identification of potential faults that can occur in a bearing during its operating life.

A considerable amount of research has been done to ascertain the capabilities of ML and DL in processing and identifying the state of health of machinery [22,23]. A convolutional neural network (CNN) model developed by Janssens et al. [24] was directly applied on frequency spectrum of the raw vibration signal. This achieved appreciable results in the absence of expert knowledge. Shao et al. [25] proposed a feature learning method for machinery fault diagnosis using autoencoders by first adopting the maximum correntropy in designing the loss function and later maximizing the main parameters of the autoencoders using the artificial fish swarm algorithm. In subsequent studies, the automatic learning ability of in-depth features by denoising autoencoders was harnessed to develop an enhanced noise reduction model for the diagnosis of faults by Meng et al. [26] through the combination of L1 and L2 regularization which enhanced the sparsity of training hyperparameters. Kong et al. [27] presented deep autoencoders (DAEs) multi-ensemble approach for bearing fault diagnosis by assembling a couple of DAEs with various activation function to extract different features during training which are grouped together and the final result is achieved by majority voting. The method proposed in [27] decreases in performance when the operating condition is not known.

Even though most of the developed DL models have achieved better performance, they are likely not to generalize well in their application in real life situations due to the few datasets that are used to train these models from scratch and also there is a great disparity between environmental and varying operating conditions. An approach being explored to deal with these issues is transfer learning which applies a model trained in a given task to perform related task in another domain [28]. Guo et al. [29] proposed a deep convolutional transfer learning network (DCTLN) which comprised condition recognition and domain adaptation modules both constructed using a 1-D CNN for machine fault diagnosis. Similarly, Zhao et al. [28] developed a multi-scale CNN (MSCNN) using a dilated convolution to achieve differential features and reduced the complexity by utilizing global average pooling. To achieve a higher diagnosis performance, Wang et al. [30] presented a multi-scale deep intra-class adaptation by modifying a pre-trained model for extracting features at a low level, and subsequently analyzing them using a multiple scale feature learner as inputs to a classifier made up of high level features.

However, to achieve a robust and high-performance model, this research considers three input formats of vibration signals and develops a diagnosis model using a transfer learning approach. In addition, this research also delves into prognostics which is essential in a complete prognostics and health monitoring (PHM) framework. Ongoing research has either focused on the diagnosis of bearing faults through classification models or estimation of remaining useful life (RUL). However, these two approaches use separate datasets, hence these two stages (diagnosis and prognostics) are conducted disparately [4,31].

To overcome this bottleneck, a deep convolutional neural network (DCNN) approach is adopted to perform diagnosis and health monitoring for bearings in rotating machinery through signal transformations and feature extraction to establish a health indicator (HI) for prognostics. The main contributions of this research work are summarized as follows:

(i) Develop a transfer learning approach by modifying the VGG16 architecture with high accuracy on ImageNet for learning low-level features in insufficient vibration datasets.

(ii) Increase overall robustness, accuracy and decrease the average computation time by transforming the raw vibration signal into spectrogram and Mel spectrogram. The Mel spectrogram representation results in the best performance.

(iii) Develop a framework that conjoins a diagnosis outcome after classification to a prognostics scheme. An overall prognostics strategy is established by constructing an HI from the fusion of significant features. The assessment to determine the likelihood of a specific statistical attribute containing deterioration data was decided by its monotonicity. The value of the HI can therefore be used to indicate the trend in degradation. This can then be fitted to a regression model for health analysis and prediction of RUL.

The rest of the paper is organized as follows. The theoretical overview of the task at hand is given in Section 2, which is followed by a description of the proposed method in Section 3. In Section 4, the experimental procedure and results are analyzed under diagnostics and prognostic frameworks. Finally, concluding remarks are presented in Section 5.

2. Theoretical background

Appropriate vibration data acquisition is vital to an effective health monitoring, fault diagnosis and prognosis of a machinery. The most popular non-destructive techniques used today are mostly focused on analysis of vibrations. In a typical rolling element bearing as shown in Fig. 1, the position of an abnormality can be determined if it corresponds with one of the frequencies given in the following formulae, calculated from the bearing's geometry:

$$f_{\text{cage}} = \frac{f_A}{2}\left[1 - \frac{D_b}{D_p}\cos\varphi\right], \tag{1}$$

$$f_{\text{outer}} = N_b \cdot f_{\text{cage}}, \tag{2}$$

$$f_{\text{Inner}} = N_b(f_A - f_{\text{cage}}), \tag{3}$$

$$f_R = \frac{D_p}{2D_b} f_A\left[1 - \left(\frac{D_b}{D_p}\right)^2 \cos^2\varphi\right], \tag{4}$$

where f_{cage}, f_{outer}, f_{inner}, and f_R represent the cage, outer race, inner race, and the roller spin frequencies respectively. N_b is the number of balls, f_A is the frequency of revolution, φ is the angle of contact, D_b and D_p are the ball and pitch diameters respectively.

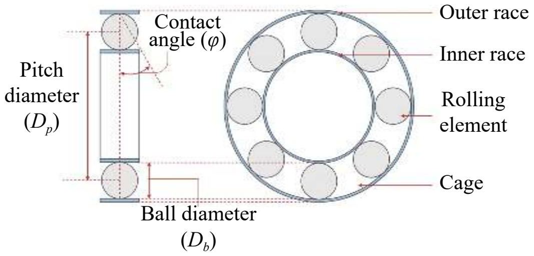


Fig. 1 Typical rolling element bearing

The impact frequencies for selected fault conditions are illustrated in Fig. 2. Based on this discernment, vibration signals in bearings are effectively classified by using a DCNN.

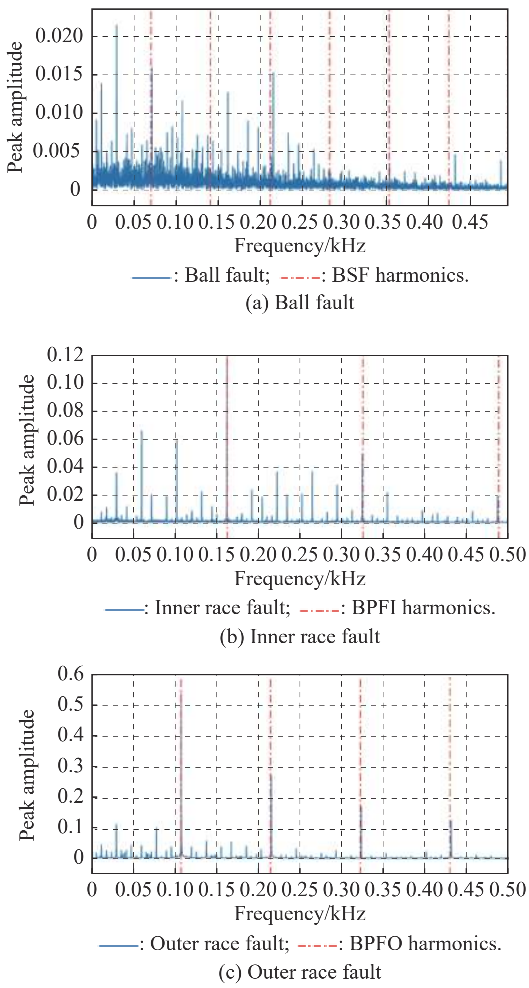


Fig. 2 Envelope spectrum of vibration signal

Transfer learning (TL) is a research exploratory field where knowledge acquired in solving a problem in a source domain is applied in a different target domain. In the area of deep learning, the TL technique results in the reduction in the training time of a model with better generalization even when large datasets are not available. TL involves the concepts of a domain D and the task T, where D consists of the feature space χ and the marginal probability $P(X)$, where $X = \{x_1, x_2, \cdots, x_n\}$. For a specific domain,

$$D = \{\chi, P(X)\}. \tag{5}$$

Conversely, a task comprises a label space γ and an objective function

$$T = \{\gamma, f(\cdot)\} \tag{6}$$

where $f(\cdot)$ can be expressed as $P(\gamma|X)$.

From the aforementioned expressions, the objective of TL is the enhancement of an objective function in the target domain by applying knowledge acquired from the source task in the source domain as represented in Fig. 3. However, there exist a difference in either the feature spaces of source and target domains ($\chi_{\text{source}} \neq \chi_{\text{target}}$) or difference in their marginal probabilities $P(X_{\text{source}}) \neq P(X_{\text{target}})$. Analogously, $\gamma_{\text{source}} \neq \gamma_{\text{target}}$ or $P(Y_{\text{source}}) \neq P(Y_{\text{target}})$, which implies differences in the source and target domain's label spaces.

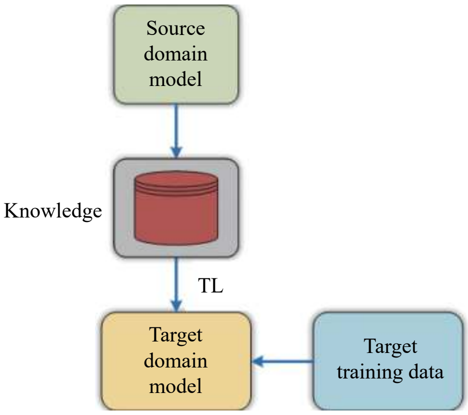


Fig. 3 Concept of TL

The VGGNet16 is a DCNN having 16 layers, hence VGG16 [32]. An advantage of this model is the convenience of loading a model pre-trained on over a million images for classifying thousand images with high accuracy.

3. Proposed method

The approach proposed in this study for health monitoring of bearing operating conditions consists of a diagnosis and prognostics section. This study adopts VGG16 architecture, capitalizing on the benefits of TL, to achieve a high-performance model by extracting distinct features from various transformations of input data. The overall model is based on the concept from transference of invariant features from a source to a target domain. The VGG16 model outperformed the other well-known pre-trained architectures such as VGG19 [32] and ResNet50 [33]. Therefore, we leverage this potential in developing a model for the purpose of classifying selected fault states. The input to the network is an RGB image of size 224×224. The subsequent blocks (conv1 to conv5) consist of convolution and pooling layers after which fully connected layers (fc6, fc7) and a SoftMax classifier serves as the final stages of the model. The modified architecture is shown in Fig. 4 where the original model is truncated after pooling layer of conv5 and weights frozen just beforethe fully connected layer. This is then flattened and fed to a new fully connected classifier. In this research, the size of fc6 and fc7 is fixed at $4\,096\times1$after several experimentations, which is followed by the final dense layer of size 10.

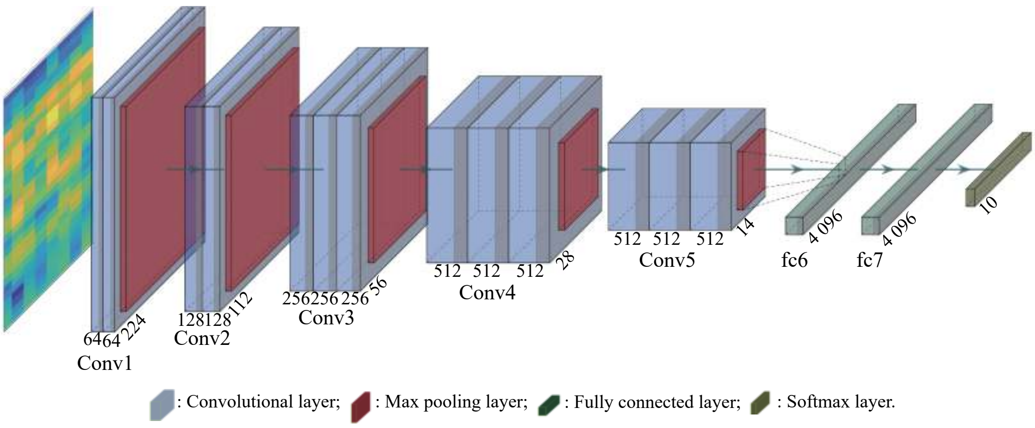


Fig. 4 Modified VGG16 architecture

4. Experimental verification

Two health management cases are conducted for fault classification and failure prognostics. The latter is designed to warn of escalating levels in vibration, based on the outcome of the former, and serves as a boundary beyond which a machinery under an identified bearing fault condition will be deemed inoperable and hence appropriate actions must be taken. The initial study groups the bearing fault data into 10 categories and the proposed model is trained on these samples. Based on the outcome of the fault diagnosis, the progression of the fault is analyzed through the extraction of significant health features and culminates into an HI for prognostics. By so doing, RUL can be estimated based on predefined thresholds to prevent catastrophic failures. The overall approach of the proposed bearing diagnosis and prognostics model is shown in Fig. 5.

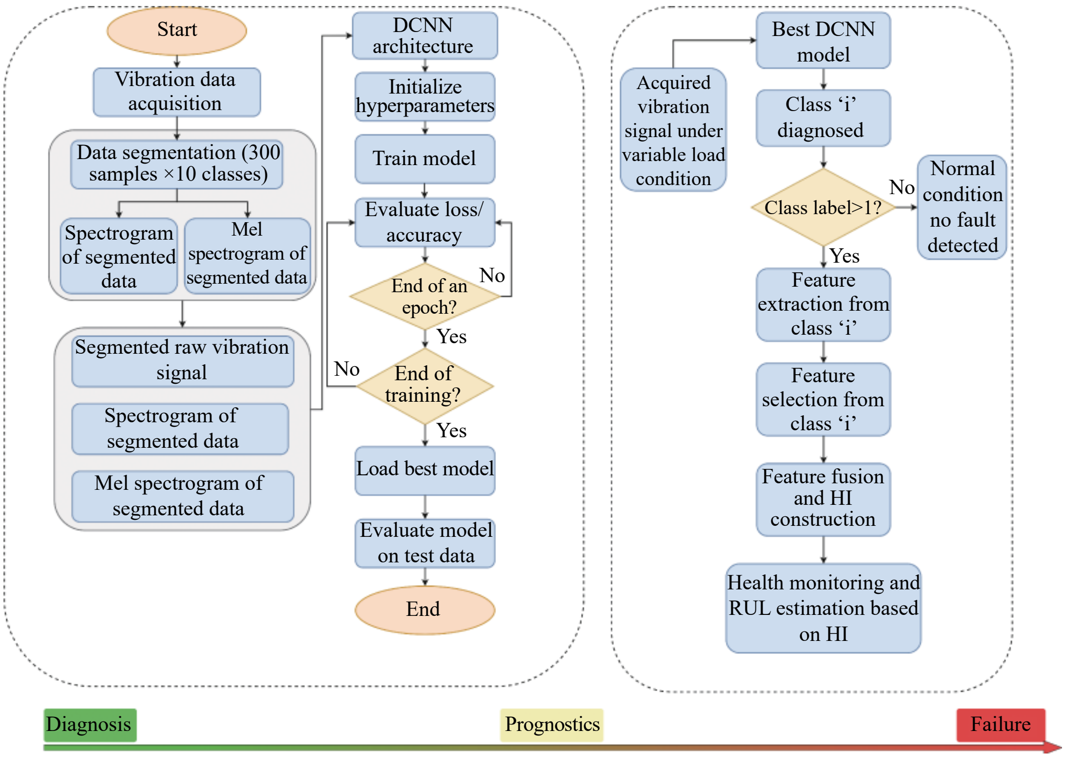


Fig. 5 Process of implementing the proposed approach

4.1 Description of experimental setup

The dataset used in the experiment is acquired from Case Western Reserve University and is collected from a 2 hp motor, a torque transducer, a dynamometer, and electronics for control. Vibration data were acquired from a single-row deep groove bearing (SKF6205-2RS) at the drive end of the motor. Further parameters of this bearing are given in Table 1.

Table 1 Parameters of bearing

Parameter	Value
Type	6205-2RS JEM SKF
Inside diameter/mm	25
Outside diameter/mm	52
Thickness/mm	15
Ball diameter/mm	8
Pitch diameter/mm	34.35
Fatigue load/kN	0.335

Single point faults were injected to the test bearing with varying fault diameters and load conditions resulting in vibration signals collected at sampling frequency (fs) of 12 kHz. These are categorized as normal condition (N), inner race fault (IR), outer race fault (OR), roller fault (ball fault). The defect frequencies of this bearing are multiples of the speed of operation which in this case varies from 1 797 rpm to 1 730 rpm at different load conditions.

4.2 Case 1: bearing fault diagnosis

4.2.1 Data segmentation

The samples used in the experimentation are obtained by segmenting the raw signals of the bearing operating conditions. The length each operating condition is given in Table 2.

Table 2 Size of vibration signal

Bearing operating condition	Length of data	
	0 hp	3 hp
Normal	240 000	480 000
Ball, inner, and outer race fault	120 000	120 000

A total of 400 data points are selected in succession without overlapping across all the individual raw data signals. An inner race fault signal with a sliding window is illustrated in Fig. 6. This process is repeated for all the operating conditions and each sample is subsequently transformed to spectrogram and Mel spectrogram. At the end of this procedure, the normal condition has 600 and 1 200 samples for 0 hp and 3 hp. The remaining fault conditions each have 300 samples for both 0 hp and 3 hp.

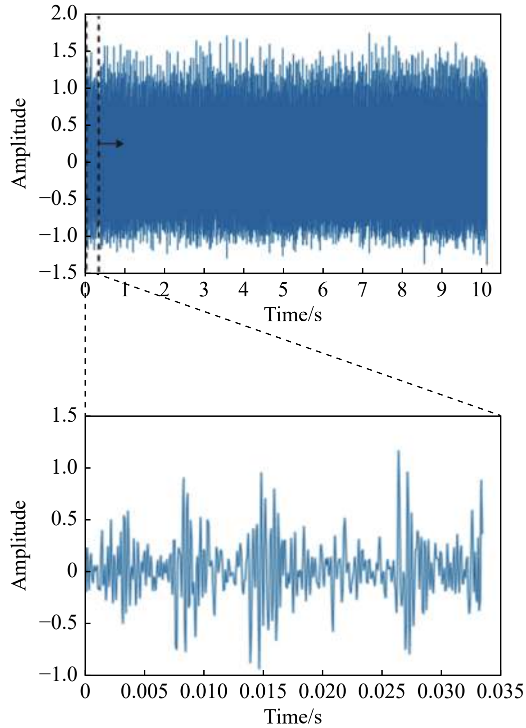


Fig. 6 Segmenting inner race fault signal

4.2.2 Data transformation

Spectral representation of signals only shows their frequency content. However, it does not expose information relating the frequency constituents at any given point in time.

Table 3 Description of bearing operating condition

Bearing operating condition	Diameter of faults/mm	Size of segment: Training 0 hp	Training 3 hp	Validation and testing 0 hp	Validation and testing 3 hp	Class label
Normal	0	480	960	120	240	1
Ball	0.18	240	240	60	60	2
Ball	0.36	240	240	60	60	3
Ball	0.54	240	240	60	60	4
Inner race	0.18	240	240	60	60	5
Inner race	0.36	240	240	60	60	6
Inner race	0.54	240	240	60	60	7
Outer race	0.18	240	240	60	60	8
Outer race	0.36	240	240	60	60	9
Outer race	0.54	240	240	60	60	10

A combined time-frequency approach can be effectual in the analysis of signal whose amplitude and frequency components vary in time.

Short time fourier transform (STFT) can be described as a method which assumes stationarity of a signal over a period. This signal is divided into a series of short (windowed) signals and the Fourier transform is performed for each of the windowed signals resulting in a complex matrix which contains both real and imaginary parts of the Fourier transform with time and frequency dimensions. The spectrogram, therefore is squared magnitude of STFT and the Mel spectrogram is the non-linear transformation of the frequency scale. A segment of vibration data and its time-frequency transforms are shown in Fig. 7.

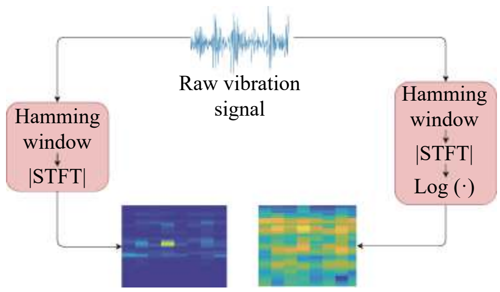


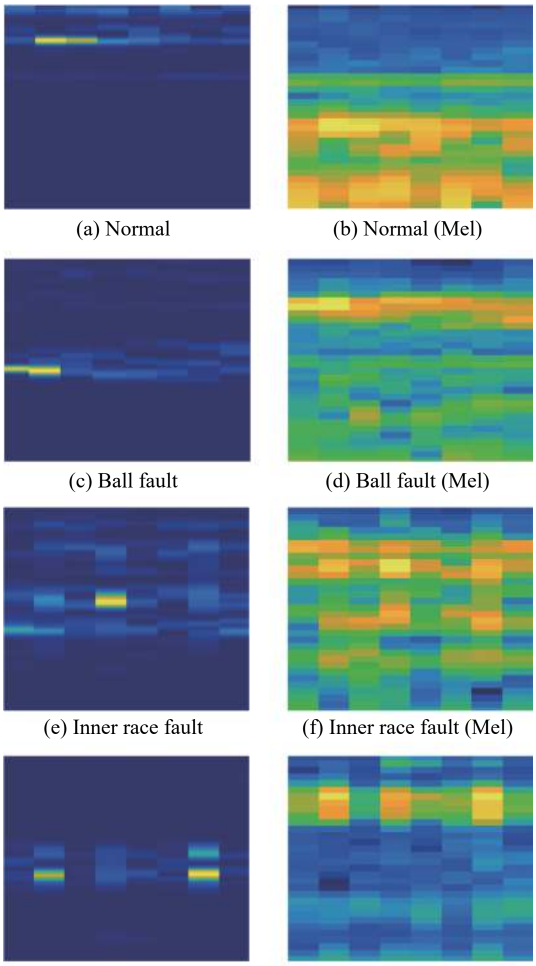
Fig. 7 Spectrogram and Mel spectrogram transformation

The parameters required for computing the spectrogram of a given signal can significantly influence the accuracy of a classifier [34]. Since there are nearly infinite combinations of such parameters, the selected values that result in best representation are given in Table 4.

Table 4 Parameters of spectrogram transformation

Parameter	Value
Sampling frequency/kHz	12
Window type	Hamming
Overlap length/%	50
Number of DFT points	256
Filter bank	Triangular
Number of filters	32
Window length	128

The x and y axes of the resultant samples represent time and frequency respectively. However, these axes are excluded from the training samples. Selected samples of spectrogram and Mel spectrogram with a 0.18 mm fault width under different operating conditions are shown in Fig. 8.



(a) Normal (b) Normal (Mel)

(c) Ball fault (d) Ball fault (Mel)

(e) Inner race fault (f) Inner race fault (Mel)

(g) Outer race fault (h) Outer race fault (Mel)

Fig. 8 Training samples of spectrogram (left) and Mel spectrogram (right) representations of operating conditions

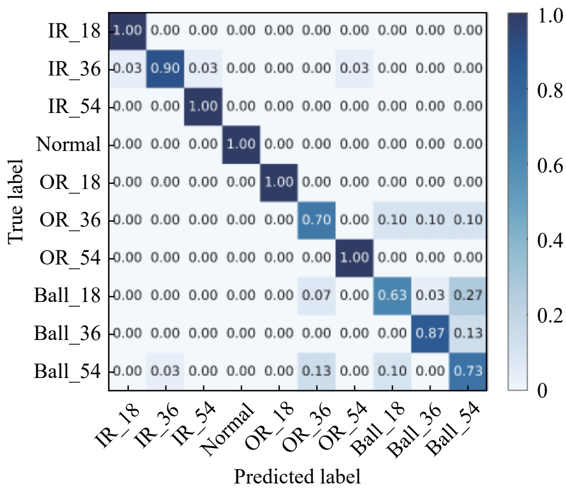
The data preprocessing procedures followed in this work are data resizing and z-score normalization. All the input images are automatically resized to 224×224 using the Keras image data preprocessing pipeline to ensure a uniform input dimension for the proposed model. After resizing the images, the pixel values (features) are normalized between the range 0 and 1 to ensure similar data distribution and faster convergence during training. Given an input image $x_{(i)}$, the normalized feature form $x_{\text{norm}(i)}$ is given by

$$x_{\text{norm}(i)} = \frac{x_{(i)} - \mu(x_{(i)})}{\sigma(x_{(i)})} \tag{7}$$

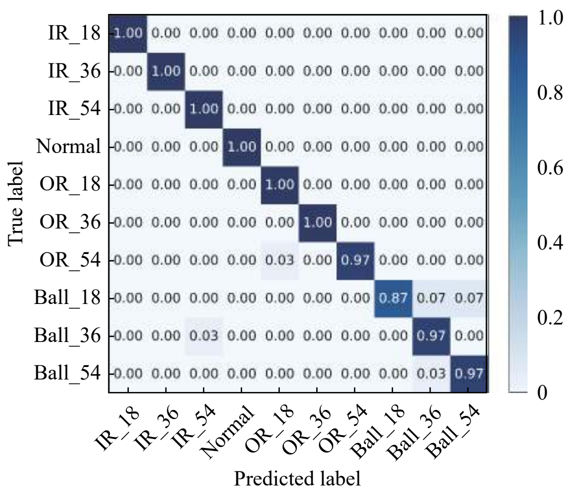
where $\mu(x_{(i)})$ and $\sigma(x_{(i)})$ are the mean and standard deviations of the image feature i.

4.2.3 Diagnosis results and analysis

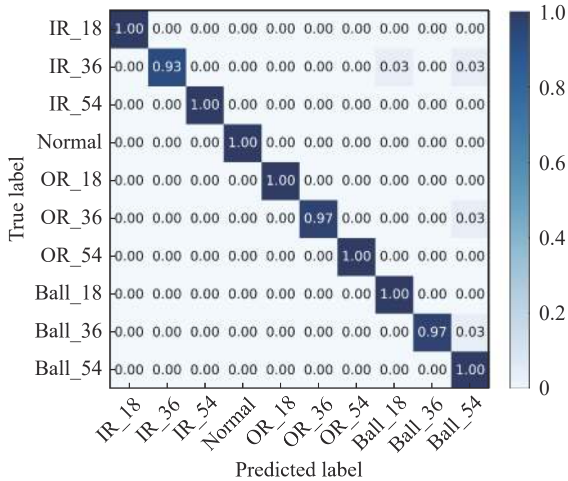
The performance of the model in classifying the ten bearing operating conditions is represented in Fig. 9 and Fig. 10. It can be observed that the model performed creditably well in most of the classification results with high precision and recall. However, ball faults seem to be the most difficult for the model to correctly classify with fault depth of 0.36 mm and 0.54 mm being the most affected across the three transforms of datasets. This can be attributed to the evidence of inner and outer race faults in the ball fault dataset. This can be seen from the ball pass frequency outer (BPFO) race which is equal to the product of the number of rolling elements and the cage rotating frequency as expressed in (2). In addition, smearing phenomenon has been observed in Fig. 12 as well as envelope spectrum of some fault depths which matches twice the ball spin frequency and other harmonics [35].



(a) Raw signal

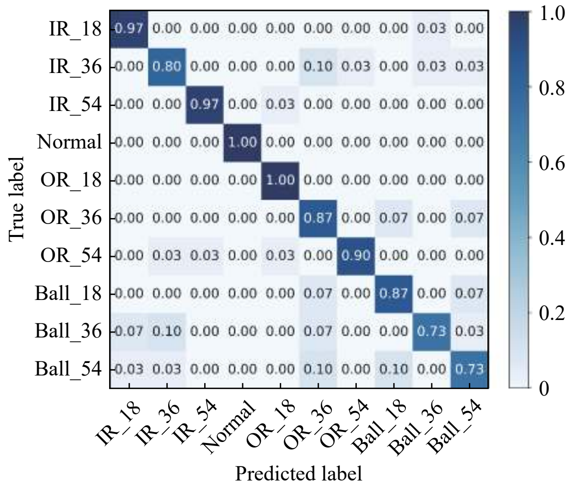


(b) Spectrogram

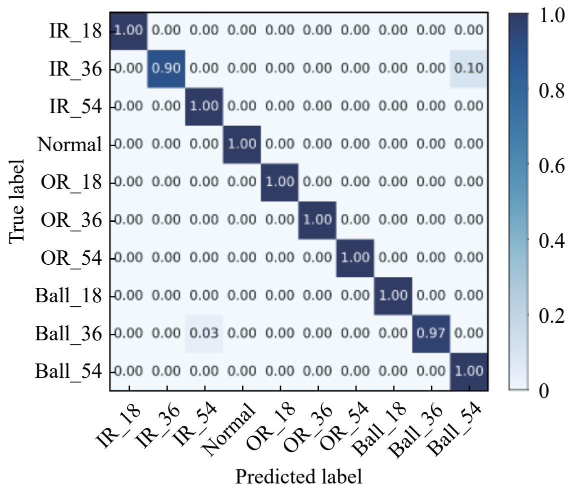


(c) Mel spectrogram

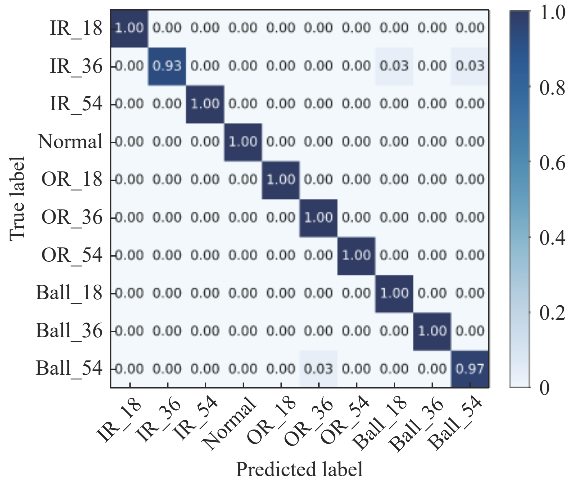
Fig. 9 Multi-class confusion matrix of the proposed method for 0 hp



(a) Raw signal



(b) Spectrogram



(c) Mel spectrogram

Fig. 10 Multi-class confusion matrix of the proposed method for 3 hp

The training and testing results after the experimental procedures are given in Table 5. The accuracy of the prediction on a test set shows good value on the Mel spectrogram data.

Table 5 Average testing accuracy and computation time comparison of the three approaches in the experiment

Load	Input data	Average accuracy/%		Average time/ (epoch/s)	Epoch for best model
		Train	Test		
0 hp	Raw signal	91.26	89.52	35	16
	Spectrogram	99.77	97.88	30	15
	Mel spectrogram	99.98	98.79	27	16
3 hp	Raw signal	98.27	91.03	41	34
	Spectrogram	100	98.97	38	34
	Mel Spectrogram	100	99.23	35	25

The effectiveness of our proposed method is compared with an ensemble deep autoencoder (DAE) [27], a feedback CNN [36], normalized CNN [37], and a multi-scale CNN-LSTM (MSCNN-LSTM) model [38]. These existing methods show good classification accuracies. However, the performance of the proposed method using Mel spectrogram data exceeds these methods as shown in Table 6.

Table 6 Comparison of average testing accuracy

Deep learning method	Average testing accuracy/%
Ensemble DAE	97.18
FCNN	98.8
Normalized CNN	98.5
MSCNN-LSTM	98.46
The proposed method	99.23

4.3 Case 2: bearing condition monitoring

In this study we consider a situation where a rotating machine operates from 0 hp to 3 hp condition as illustrated in Fig. 11 under the assumption that only one fault has occurred at a given time. In addition, analogous to Fig. 7 the signal is demarcated into 300 segments with each segment having 400 data points.

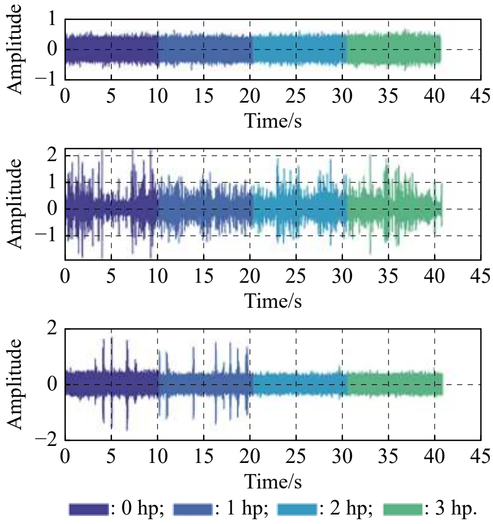


Fig. 11 Vibration signal at different loads for ball fault of width 0.18, 0.36, and 0.54 mm from top to bottom respectively

4.3.1 Degradation assessment

Fault diagnosis depends mainly on extracting a set of features from sensor data that can distinguish between fault classes of interest and detect and isolate a particular fault at its early initiation stages. Since the sensor data can be noisy, it is vital for an effective signal processing approach. In this study the vibration signal is analyzed at the various fault depths and its impact of increasing load from 0 hp to 3 hp condition. A sample of this is shown in Fig. 12 for the impact of increasing load on a ball fault.

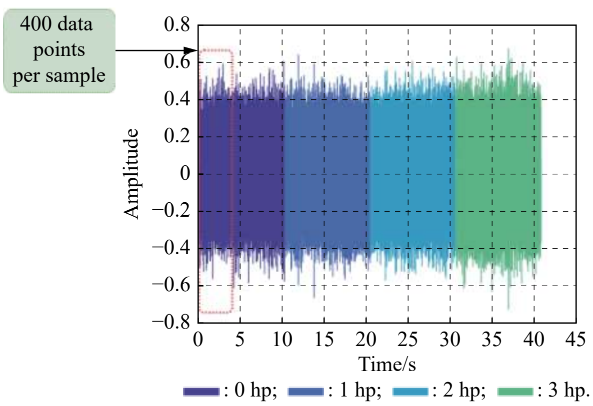
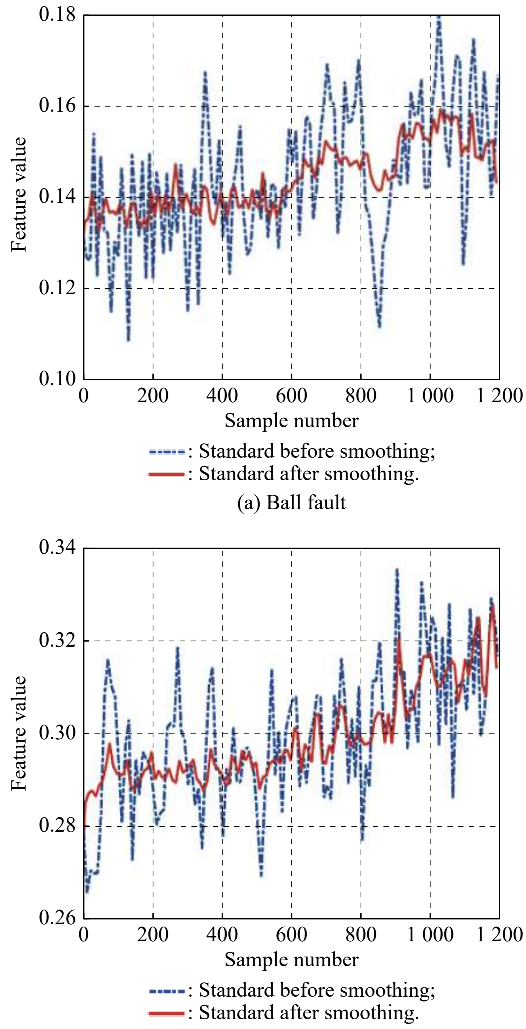


Fig. 12 Degradation in ball fault (0.18 mm) over varying load conditions

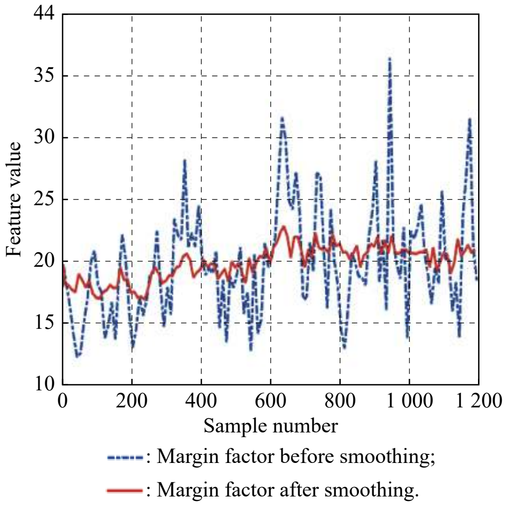
4.3.2 Health indicator construction

The process of constructing the HI begins with the acquisition of relevant data. The time-domain representation of the signal shown in Fig. 12 indicates an increasing trend in amplitude as load increases. Subsequently, some time and frequency domain features are extracted to analyze the health condition information. The following time domain statistical features are extracted: mean, standard deviation, skewness, kurtosis, peak2peak, root mean square (RMS), crest factor, shape factor, impulse factor, margin factor, energy. Conversely, the spectral kurtosis of the mean, standard deviation, skewness, and kurtosis are extracted from the signal in frequency domain. The HI is developed from the aforementioned time-domain and frequency domain extracted features for the purpose of prognostics. The extracted features are preprocessed to remove noise which can interfere in prognostic implementation. The effect of filtering out the noise using a mean filter with a lag window of five steps is shown in Fig. 13 for selected domain representations.



(a) Ball fault

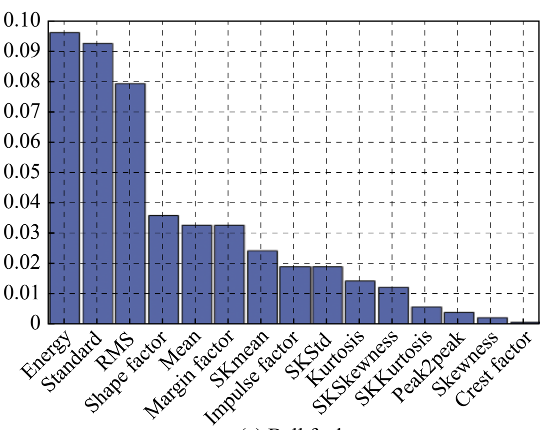
(b) Inner race fault



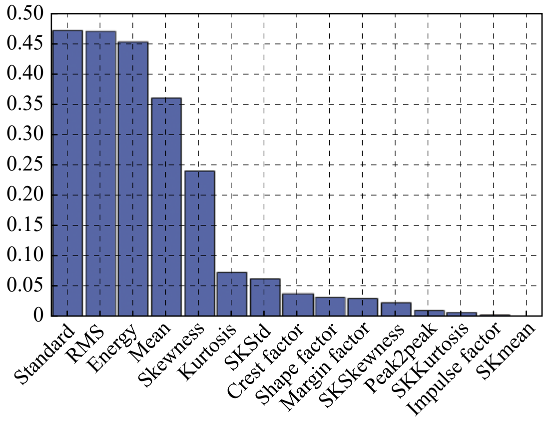
(c) Outer race fault

Fig. 13 Selected significant feature process

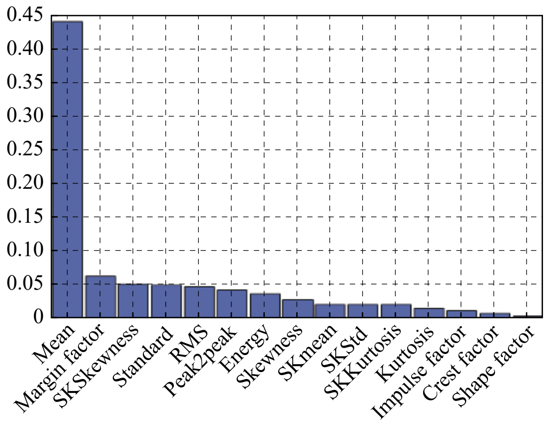
Machines and equipment failure processes are irreversible in practical applications. This implies that a defective part cannot recover by itself without a certain form of maintenance or intervention. An acceptable HI must have a monotonic raising or lowering pattern to correspond with the irrecoverable degradation processes. This effect is known as monotonicity and it is used in quantifying the features by merit [4]. This is an intrinsic characteristic of an HI itself, without considering its interactions with certain other variables and often described using a formula based on the HI pattern. Ranking the extracted and smoothed features according to importance as shown in Fig. 14 is essential for further processing. The importance of features is obtained by taking the average of the difference between positive and negative derivatives for each feature. Features above a predefined threshold are selected for later feature fusion.



(a) Ball fault



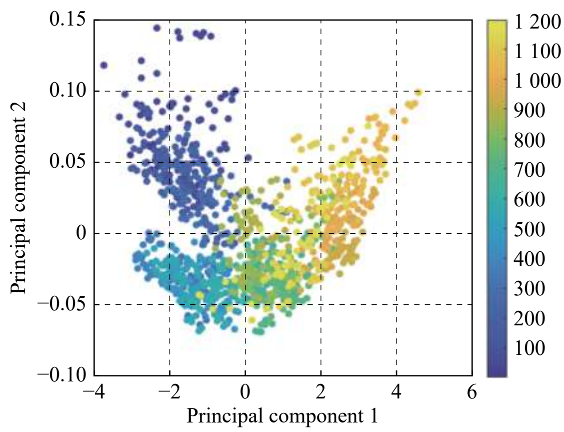
(b) Inner race fault



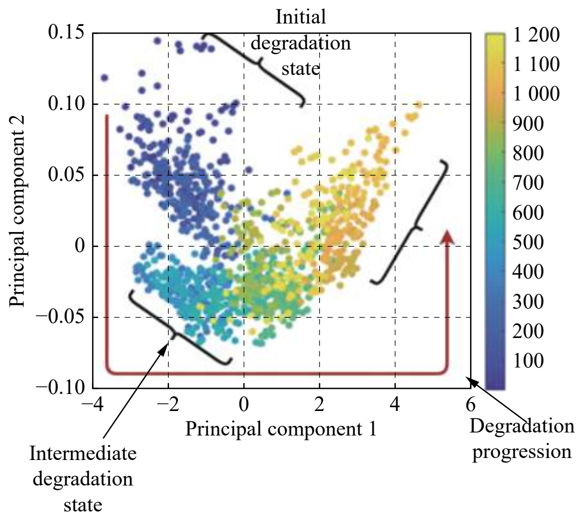
(c) Outer race fault

Fig. 14 Monotonicity of extracted statistical features

Finally, an unsupervised linear transformation approach is adopted to determine unique and appropriate features that represent the relevant information about the bearing's condition degeneration. The principal component analysis (PCA) transforms the original data, which comprises the selected features, to a new lower-dimensional subspace of orthogonal data referred to as principal component (PC). Patterns in the data are identified as a result of the correlation between features. An example of a ball fault is given in Fig. 15.



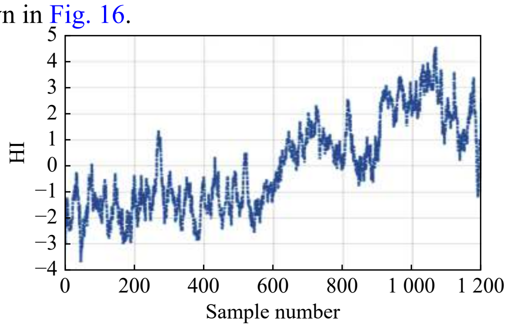
(a) PCA of a ball fault



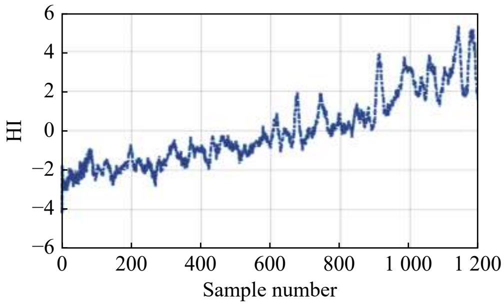
(b) Degradation states in a ball fault

Fig. 15 Visualization of different degradation states of a ball fault

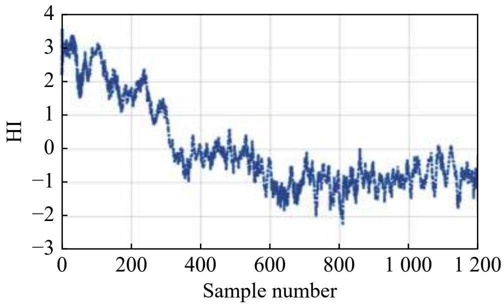
From experimental simulations, it is observed that PC1 has the best HI due to the progression in the data as shown in Fig. 16.



(a) Ball fault



(b) Inner race fault



(c) Outer race fault

Fig. 16 HI for prognostics

4.3.3 Prediction of degradation behavior based on feature fusion

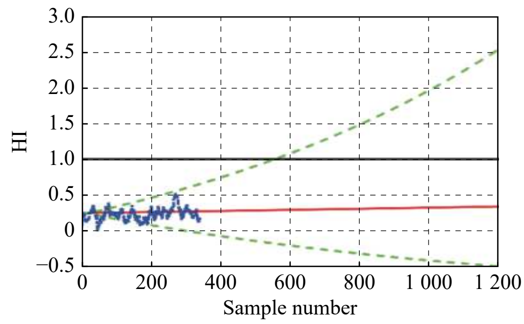
A data-driven prognostic methodology based on the systematic prediction is proposed in this section. A model based on exponential deterioration is fit to the HI obtained in the previous section. The degradation model adopted in this study for estimating the state of health until a predetermined threshold is reached consists of stochastic and deterministic components. The HI at a given sampling instance k is modeled as

$$\mathrm{HI}(k)=\theta\exp\left(\beta(k)+\varepsilon-\frac{\sigma^2}{2}\right)+\phi \tag{8}$$

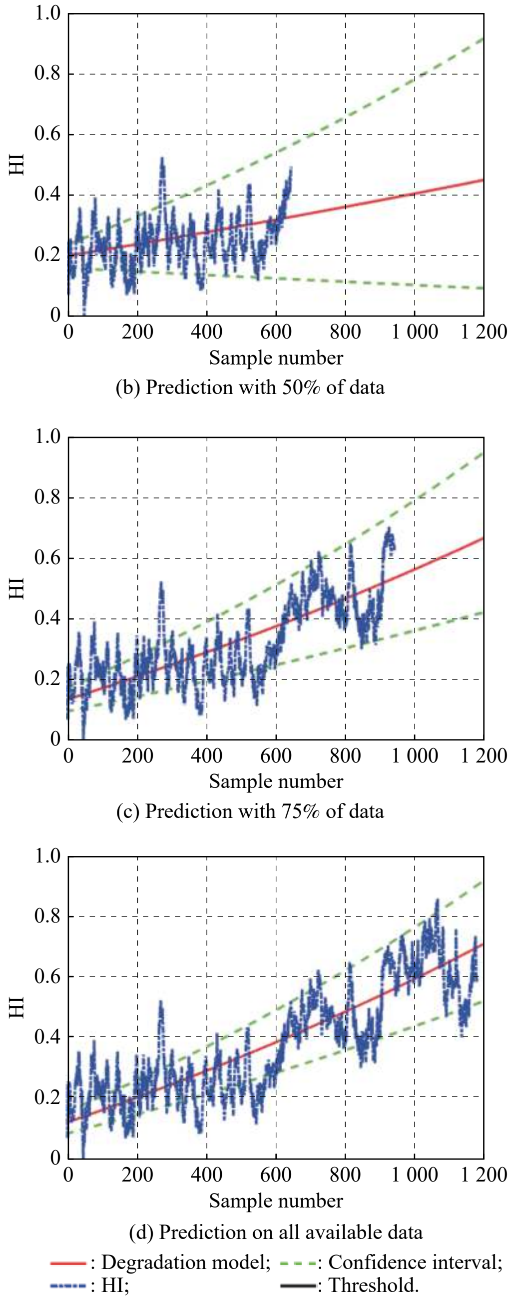
where θ is a lognormal distributed variable and β is a Gaussian distributed random variable representing the stochastic components are random variables, ε is a Gaussian white noise with zero mean and variance σ^2. ϕ is a deterministic intercept term which is constant.

Threshold determination is typically based on an equipment historical record or other domain-specific information [39]. The HI is normalized to range from 0 and 1, and the RUL is estimated to be the time from the current HI till the deterioration reaches a threshold. For the purposes of maintenance scheduling, two levels of threshold alerts can be set. A warning level to indicate that the HI has reached or exceeded at least 0.8 and prompt the scheduling of maintenance task, and a critical level of 1.0, beyond which maintenance has to be performed. Using Matlab functions, the exponential degradation model is developed with relevant parameters. The model predicts and updates in real time by estimating future loading of the bearings and detecting slopes at various instances, thereby ignoring past observations and re-initializing estimation based on past information. This results in a model that makes predictions based on expected loads.

The degradation process is represented in Fig. 17, where there is no evidence of degeneration preceding the 400th sample. This gives rise to an RUL which is potentially immeasurable as the slope is close to zero. Subsequently, degradation becomes evident as the load increases close to the 600th sample. Beyond this time, the model can be used to estimate the RUL with higher confidence as more data becomes available.



(a) Prediction with 25% of data



(b) Prediction with 50% of data

(c) Prediction with 75% of data

(d) Prediction on all available data

Fig. 17 Degradation evolution and prediction results

5. Conclusions

This paper proposes a health monitoring approach for rolling bearings comprising a diagnosis and prognostics framework. In developing a robust model for the classification of bearing operating conditions, a method based on TL is adopted by transferring the pre-trained DCNN model to three transforms of vibration signals under varying operating conditions. Subsequent comparative experiments were carried out to analyze the performance of these three signals transforms under two operating conditions. It was realized that Mel spectrogram transformed signal offered the best classification results owing to the ability of the model to extract distinguishing features.

Furthermore, this paper explores the health monitoring and life prediction of a faulty bearing by generating an HI by fusing significant statistical features. The assessment to determine the likelihood of a specific statistical attribute subsuming deterioration was decided by its monotonicity. The value of the fused HI by principal component analysis can therefore be used to indicate the trend in degradation. It has been shown that the features extracted can provide an advanced indication of rising bearing defects under a given bearing fault condition and enhance bearing deterioration assessment. Developing deep learning approaches for fault diagnosis and prognostics is very important. Therefore, in realistic industrial scenarios, the proposed approach intends to increase equipment reliability.

Nevertheless, due to large variability in the HI, degradation tracking and RUL estimation becomes a challenge as prediction accuracy is highly dependent on the efficiency of an HI. This is a problem to be resolved. In the future, the authors will further investigate this issue by implementing different deep learning approaches to achieve better prognostics performance.

References

[1] HARRIS T A. Rolling bearing analysis. New York: John Wiley & Sons, 2001.

[2] ISLAM M M, KIM J M. Reliable multiple combined fault diagnosis of bearings using heterogeneous feature models and multiclass support vector machines. Reliability & Engineering Systems Safety, 2019, 184: 55–66.

[3] SAMANTA B, AL-BALUSHI K R, AL-ARAIMI S A. Artificial neural networks and genetic algorithm for bearing fault detection. Soft Computing, 2006, 10(3): 264–271.

[4] LEI Y G, LI N P, GUO L, et al. Machinery health prognostics: a systematic review from data acquisition to RUL prediction. Mechanical Systems and Signal Processing, 2018, 104: 799–834.

[5] AHMED H, NANDI A K. Condition monitoring with vibration signals: compressive sampling and learning algorithms for rotating machines. Hoboken: John Wiley & Sons, 2019.

[6] MCCORMICK A C, NANDI A K. A comparison of artificial neural networks and other statistical methods for rotating machine condition classification. IEE Colloquium on Modeling and Signal Processing for Fault Diagnosis, 1996. DOI:10.1049/ic:19961372.

[7] PRIYA V S, MAHALAKSHMI P, NAIDU V P S. Bearing health condition monitoring: wavelet decomposition. Indian Journal of Science and Technology, 2015, 8(26): 1–7.

[8] NAYANA B R, GEETHANJALI P. Analysis of statistical time-domain features effectiveness in identification of bear-

ing faults from vibration signal. IEEE Sensors Journal, 2017, 17(17): 5618–5625.

[9] MCFADDEN P D, TOOZHY M M. Application of synchronous averaging to vibration monitoring of rolling element bearings. Mechanical Systems and Signal Processing, 2000, 14(6): 891–906.

[10] AYAZ E. Autoregressive modeling approach of vibration data for bearing fault diagnosis in electric motors. Journal of Vibroengineering, 2014, 16(5): 2130–2138.

[11] LI X H, FU P. Rolling bearing fault diagnosis based on one-dimension blind source separation. Journal of Electronic Measurement and Instrumentation, 2014, 27(6): 535–542. (in Chinese)

[12] PATIDAR S, SONI P. An overview on vibration analysis techniques for the diagnosis of rolling element bearing faults. International Journal of Engineering Trends and Technology, 2013, 4(5): 1804–1809.

[13] TIAN J, MORILLO C, AZARIAN M H, et al. Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with *k*-nearest neighbor distance analysis. IEEE Trans. on Industrial Electronics, 2015, 63(3): 1793–1803.

[14] WANG Y X, XIANG J W, MARKERT R, et al. Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: a review with applications. Mechanical Systems and Signal Processing, 2016, 66: 679–698.

[15] TSAO W C, LI Y F, DU LE D, et al. An insight concept to select appropriate IMFs for envelope analysis of bearing fault diagnosis. Measurement, 2012, 45(6): 1489–1498.

[16] BRANDT A. Noise and vibration analysis: signal analysis and experimental procedures. New York: John Wiley & Sons, 2011.

[17] LIU H M, LI L F, MA J. Rolling bearing fault diagnosis based on STFT-deep learning and sound signals. Shock and Vibration, 2016: 1–12.

[18] KANKAR P K, SHARMA S C, HARSHA S P. Fault diagnosis of ball bearings using continuous wavelet transform. Applied Soft Computing, 2011, 11(2): 2300–2312.

[19] DYBAŁA J, ZIMROZ R. Rolling bearing diagnosing method based on empirical mode decomposition of machine vibration signal. Applied Acoustics, 2014, 77: 195–203.

[20] PANG B, TANG G J, TIAN T, et al. Rolling bearing fault diagnosis based on an improved HTT transform. Sensors, 2018, 18(4): 1203.

[21] WEI Z X, WANG Y X, HE S L, et al. A novel intelligent method for bearing fault diagnosis based on affinity propagation clustering and adaptive feature selection. Knowledge-Based Systems, 2017, 116: 1–12.

[22] LIU R N, YANG B Y, ZIO E, et al. Artificial intelligence for fault diagnosis of rotating machinery: a review. Mechanical Systems and Signal Processing, 2018, 108: 33–47.

[23] ZHANG S, ZHANG S B, WANG B N, et al. Deep learning algorithms for bearing fault diagnosticsx –a comprehensive review. IEEE Access, 2020, 8: 29857–29881.

[24] JANSSENS O, SLAVKOVIKJ V, VERVISCH B, et al. Convolutional neural network based fault detection for rotating machinery. Journal of Sound and Vibration, 2016, 377: 331–345.

[25] SHAO H D, JIANG H K, ZHAO H W, et al. A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. Mechanical Systems and Signal Processing, 2017, 95: 187–204.

[26] MENG Z, ZHAN X Y, LI J, et al. An enhancement denoising autoencoder for rolling bearing fault diagnosis. Measurement, 2018, 130: 448–454.

[27] KONG X G, MAO G, WANG Q B, et al. A multi-ensemble method based on deep auto-encoders for fault diagnosis of rolling bearings. Measurement, 2020, 151: 107132.

[28] ZHAO B, ZHANG X M, ZHAN Z H, et al. Deep multi-scale convolutional transfer learning network: a novel method for intelligent fault diagnosis of rolling bearings under variable working conditions and domains. Neurocomputing, 2020, 407: 24–38.

[29] GUO L, LEI Y G, XING S B, et al. Deep convolutional transfer learning network: a new method for intelligent fault diagnosis of machines with unlabeled data. IEEE Trans. on Industrial Electronics, 2019, 66(9): 7316–7325.

[30] WANG X, SHEN C Q, XIA M, et al. Multi-scale deep intra-class transfer learning for bearing fault diagnosis. Reliability Engineering & System Safety, 2020, 202: 107050.

[31] KHAN S, YAIRI T. A review on the application of deep learning in system health management. Mechanical Systems and Signal Processing, 2018, 107: 241–265.

[32] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition. Proc. of the 3rd International Conference on Learning Representations, 2015. DOI: arXiv:1409.1556.

[33] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition. Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770–778.

[34] NYKAZA E T, NETCHAEV A, BUNKLEY S, et al. Objectively choosing spectrogram parameters to classify environmental noises. INTER-NOISE and NOISE-CON Congress and Conference Proceedings, 2016, 253(1): 7336–7343.

[35] SMITH W A, RANDALL R B. Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study. Mechanical Systems and Signal Processing, 2015, 64: 100–131.

[36] AN F-P. Rolling bearing fault diagnosis algorithm based on FMCNN-sparse representation. IEEE Access, 2019, 7: 102249–102263.

[37] ZHAO B, ZHANG X M, LI H, et al. Intelligent fault diagnosis of rolling bearings based on normalized CNN considering data imbalance and variable working conditions. Knowledge-Based Systems, 2020, 199: 105971.

[38] CHEN X H, ZHANG B K, GAO D. Bearing fault diagnosis base on multi-scale CNN and LSTM model. Journal of Intelligent Manufacturing, 2020: 1–17.

[39] BLOCH H P, GEITNER F K. Machinery failure analysis and troubleshooting. Waltham: Elsevier, 2012.

Biographies

KULEVOME Delanyo Kwame Bensah was born in 1983. He received his B.Eng. degree from Accra Institute of Technology and M.Eng. degree in electronic science and technology in 2019 from University of Electronic Science and Technology of China, Chengdu, China, where he is currently pursuing his Ph.D. degree in information and communication engineering. His research interests include prognostics and health management of systems, signal processing, and deep learning.
E-mail: kdelanyo@ieee.org

WANG Hong was born in 1974. He received his B.S., M.S., and Ph.D. degrees from Northwestern Polytechnical University, Chongqing University, and University of Electronic Science and Technology of China (UESTC), respectively. He has been a faculty member with UESTC since 2003. From 2007 to 2009 he was engaged in doctoral research with the Second Research Institute of Civil Aviation Administration. From 2009 to 2010 he was a research scholar with Polytechnic Institute of New York University and a research assistant with New Jersey Institute of Technology, USA. His present areas of interest include radar signal processing, avionics, aeronautical telecommunication, and surveillance technologies in ATC.
E-mail: hongw@uestc.edu.cn

WANG Xuegang was born in 1962. He received his Ph.D. degree from Xidian University in 1992. He is now a professor and Ph.D. supervisor with University of Electronic Science and Technology of China. His research interests include radar signal processing, and millimeter wave radar.
E-mail: xgwang@uestc.edu.cn