BJES Basrah Journal for Engineering Sciences

Basrah Journal for Engineering Sciences, Vol. 25, No. 1, (2025), 32-41

Original Article

Journal homepage: www.bjes.edu.iq
ISSN (Online): 23118385, ISSN (Print): 18146120

Developing an Efficient Technique for Predicting Ball Bearing Defects Based on RDE Method Using CNN

Haidar A. Alhajjaj [1]*, Jaafar K. Alsalaet [2]

[1,2] *Department of Mechanical Engineering, College of Engineering, University of Basrah, Basrah, Iraq*
E-mail addresses: haidaradnan08@gmail.com , jaafar.ali@uobasrah.edu.iq

Article Info

Article history:
Received: 22 September 2023
Revised: 10 October 2023
Accepted: 7 January 2024
Published: 16 August 2025

Keywords:
Convolutional Neural Network (CNN), Ball bearing faults, Reverse dispersion entropy (RDE), RDEgram, Prediction methods.

https://doi.org/10.33971/bjes.25.1.5

Abstract

The accurate prediction of machinery faults is considered an effective strategy to increase the operation life of machines, ensure smooth operation, and provide a safe environment. Accordingly, the demands on predictive tools such as machine learning to detect machinery faults before catastrophic failure occurs has increased rapidly. In this research, a diagnosis algorithm based on using a 2D color-coded map as the input to a deep artificial neural network is proposed. These maps are called RDEgram after the processing of vibrational signals based on reverse dispersion entropy (RDE) method. The effectiveness of the proposed algorithm is investigated by testing its capability to detect different faults located at different locations on ball bearings under constant speed conditions. First, the squared envelope signal is extracted by applying the short time Fourier transform to vibration signal. Then, the RDE is used to process the squared envelope to detect the range of frequencies at which the transients occur. The RDEgram color-coded map is used to represent the RDE values as a function of frequency and frequency resolution. The maps from different fault features are collected to form the diagnostic patterns. Finally, a pretrained convolutional neural network (CNN) is applied to learn the feature pattern and diagnose the bearing faults. The CNN is trained using fixed-speed data and then it is applied to diagnose faults in the test data recorded at the same speed. The prediction method adopted in the current research shows a 100% level of accuracy for predicting two types of faults (pit and slot) located at various positions a ball bearing (KOYO 1205 C3 type) running at two constant speeds (25 and 30 Hz).

1. Introduction

Ball bearings are essential components for carrying the load-elements in rotating machines, and their faults can cause high risk to the entire machine. Accordingly, it has been found that machine learning can increase the life of machines and provide save environment to the laborers while working.

Defects in bearings usually appear in the form of vibration signals as disturbances in the time and frequency domains. The disturbance occurs as a result of forces produced when some defects show up in the ball or the race of bearings such as pits, scratches or slots [1]. Vibration signals carry the dynamics of malfunctions in rotating machinery; therefore, they can be employed to detect faults [1-3]. Every rotating machine has a unique vibration signature. The aspects of measured vibration signals have information regarding the machine's condition and most of the mechanical faults [4, 5]. Consequently, several efficient methods have been suggested for finding the diagnostic features (DFs) related with the aspects of vibration signals that helps diagnosing accurately faults such as the adaptive feature extraction method for bearing fault diagnosis [6]. Diagnosis based on artificial intelligence induces advances in the field of machinery failure detection technology, which can efficiently analyze the collected data and provide reliable diagnostic results automatically [7]. Among the various intelligence methods, deep learning (DL) has attracted much attention and been used in a variety of fields [8]. Convolutional neural networks (CNNs) are a division of the deep learning and have made it possible to recognize the patterns in large inputs such as raw signals or images by extracting characteristic features. In this regard, Zhang et al. [9] adopted CNN to train raw 1-D temporal vibration signal data as the input into a 2D image for bearing faults diagnosis. Zhang et al [10] found a new technique for training a CNN and used it for diagnosing bearing faults. The raw and noisy vibration signals under different loading conditions were used as the inputs of the CNN in order to show the robustness and high accuracy of the proposed technique. The features, which were extracted by the CNN from the signals, were also visualized and explained. In a similar study, Huang et al. used deep decoupling CNN to detect compound faults [11]. Wang et al. [12] used the images produced from multiple sensor fusion that map different fault features into the CNN to diagnose faults in the wind power rig. In Hoang and Kang's [13], gray-scale vibration images were extracted from 1-D vibration signals and used as inputs to the CNN in fault diagnosis of rolling element bearing. Udmale et al. used the kurtogram images as the input to a CNN to facilitate the task of identifying the dominant features that represent the bearing faults [14]. Another study conducted by Udmale et al included using the DL sequence models to analyze the kurtogram sequential data to eliminate the feature selection exercise in the diagnosis of rolling element bearing faults [15]. Jiang et al. proposed a method named dual attention dense convolutional network to handle some deficiencies of the CNN [16]. Azamfar et al. applied transfer learning (TL) with 1D CNN model to diagnose gearbox faults [17]. Cao et al. proposed the temporal convolutional network with residual self-attention mechanism to predict the remaining life of roller bearings. Their method can learn both time-frequency and temporal information of signals [18].

In this study, the measured vibration signals are processed using reverse dispersion entropy (RDE) method. Then, colored-coded maps are extracted from the RDE and used as an input to the convolutional neural network (CNN). The main goal of the current research is to diagnose accurately two types of faults (pit and slot) located at various positions (inner race, outer race, and ball element) of a ball bearing (KOYO 1205 C3 type) running at constant speeds (25 Hz and 30 Hz).

2. Theoretical overview

2.1. Vibration signal properties

Due to the wide-range applications of transients or non-stationary vibration signals, a lot of researchers characterize these vibration signals and study deeply their behavior, for instance Al-Raheem [19] and Antoni and Randall [20]. These signals can be described using the following equation:

$$x\,(t) = \sum_{r} A_r\, S\,(t - T_r) + n(t) \tag{1}$$

Where:

A_r : represents the amplitude that varies with time due to the fluctuation of the applied force.

T_r : is the impacting time of rth order ($T_r = r\,T + \tau_r$),

T : stands for the average fault time.

τ_r : is the fluctuation that occurs during the fault time owning to the variation in speed or slippage in bearings.

$S(t)$: refers to the waveform initiated by non-stationary events which describe the decaying ringing, and $n(t)$ identifies the additive background noise. $S(t)$ can be estimated as follows.

$$\mathrm{S}(t) = \mathrm{e}^{-\xi\omega_{\mathrm{n}}\mathrm{t}} \cos\left(\sqrt{1-\xi^2}\;\omega_{\mathrm{n}}t + \varphi\right) \tag{2}$$

Where:

ξ : indicates to the damping ratio, ω_n : refers to the natural or ringing frequency, and φ : represents the phase angle or shifting angle as a response to the excited force.

2.2. Entropy

The entropy can be defined as $H = -\sum_i p_i \, ln\, p_i$ which is usually used in the thermodynamics and signal processing sciences. In signal processing science, entropy was first defined by Shannon as an irregularity or uncertainty measurement and hence refers to the amount of useful information. A higher entropy indicates that the amount of irregularity or uncertainty is high while the amount of information is low. Conversely, if the entropy value is small, the information is high, whereas the level of irregularity or uncertainty is low. The stationary signal could be described as any system in an equilibrium state. If the entropy level is low, this gives a sign that there is a transient event or non-stationarity in the vibration signal. In addition to the signal processing field, entropy has been employed in other fields including thermodynamics, physics, and economy [21]. Even though Shannon entropy can successfully determine the amount of information in a signal, it does not tell anything about the organization of the information and the relation between the entropy and time.

To study the rate of change of entropy in any process, some generalizations of Shannon's theorem are introduced such as Kolmogorov–Sinai entropy and Rényi entropy [21]. However, the applications of these techniques to measure the complexity of limited length and noisy signals were not so successful. To tackle the deficiencies of some of the complexity measurement algorithms, approximate entropy (ApEn) and sample entropy (SamEn) were introduced [22-24]. (ApEn) was formulated with the same philosophy of Kolmogorov-Sinai entropy but with improved performance. Low (ApEn) value indicates that the system is persistent, repetitive, and predictive. Richman and Moorman [23] have found that (ApEn) is biased and heavily affected by data length and the results are not consistent due to dependency on filtering level. Therefore, they introduced sample entropy (SampEn) as an alternative to (ApEn). Although (SampEn) is powerful, it is computationally demanding especially for long signals [25]. Another regularity indicator that is conceptually simple and computationally fast is the PE [26]. However, PE considers the ordinal structure of the signal and does not take into account the mean value of the amplitudes, nor the differences between the values. Recently, the dispersion entropy [25], which will be discussed soon, is introduced as an alternative to PE.

2.3. The Infogram

One of the most serious limitations of the kurtogram is that kurtosis decreases as the repetition rate of transients increases [22]. An example is when the transients become too frequent such that they overlap each other. The other shortage of the kurtogram is that when the transients are so spaced such that there is one impulse event in the entire record, in which case the kurtosis is maximal. This is the reason why the kurtogram is highly sensitive to impulsive noise. For these (and other) reasons, Antoni proposed the Infogram as an alternative to the kurtogram.

Given a time series signal $x(n)$, the SE, $u(n; f_k, \Delta f)$, is the instantaneous energy flow, at a given instant $n\Delta t$, frequency f_k, and frequency resolution Δf which is the amplitude squared of the complex envelope [22]:

$$u\left(n; f_k, \Delta f\right) = \left| \sum_{i=nR}^{N_w + nR} x(i)\; w(i - nR) e^{-j2\pi ki/N_w} \right|^2 \tag{3}$$

Where: N_w is the processing block size or STFT window size, $\Delta f = Fs/N_w$, $f_k = k\,\Delta f$, $w(n)$ is the window function, and R determines overlapping percentage between successive blocks.

Despite the envelope can be evaluated by different techniques such as FIR filter of fast kurtogram [24], STFT imposes flexibility in the selection of window size and, hence, frequency resolution. With the development of very efficient algorithms to calculate the STFT such as the fastest Fourier transform in the West (FFTW), the calculations complexity is not a matter of concern. Antoni [22] proposed definition of the entropy by interpreting the squared instantaneous energy flow, normalized by its average, as probability distribution. To obtain analogy with SK, Antoni proposed using negative of the entropy or negentropy. The spectral negentropy $\Delta I_u\,(f, \Delta f)$ of the squared envelope is given by:

$$\Delta I_u(f, \Delta f) = \left\langle \frac{u(n; f, \Delta f)^2}{\left\langle u(n; f, \Delta f)^2 \right\rangle}\, ln\left(\frac{u(n; f, \Delta f)^2}{\left\langle u(n; f, \Delta f)^2\right\rangle}\right) \right\rangle \tag{4}$$

Where:

$\langle u(n; f, \Delta f)^2 \rangle = \frac{1}{\mathrm{L}} \sum_{n=0}^{L-1} u(n; f, \Delta f)^2$ is the average of the squared energy. Similarly, spectral negentropy in the frequency domain can be defined using SES.

$U(\alpha; f_k, \Delta f)$, which is the Fourier transform of $u(n; f_k, \Delta f)$

$$\Delta I_U(f, \Delta f) = \left\langle \frac{U(n; f, \Delta f)^2}{\langle U(\mathrm{n}; f, \Delta f)^2 \rangle} ln \left(\frac{U(n; f, \Delta f)^2}{\langle U(\mathrm{n}; f, \Delta f)^2 \rangle} \right) \right\rangle \tag{5}$$

The maximum value of SE negentropy $\Delta I_u(f, \Delta f)$ corresponds to an impulse in the SE while minimum value corresponds to a sinusoidal SE (i.e., cyclostationary signal). This is the inverse behavior of the SES negentropy $\Delta I_U(f, \Delta f)$. Therefore, both $\Delta I_u(f, \Delta f)$ and $\Delta I_U(f, \Delta f)$ are used jointly in order to detect energy fluctuation in different types of signals. Hence, the average negentropy $\Delta I_\rho(f, \Delta f)$ is proposed by Antoni [22]:

$$\Delta I_\rho(f, \Delta f) = \rho \Delta I_u(f, \Delta f) + (1 - \rho) \Delta I_U(f, \Delta f) \tag{6}$$

where ρ is taken as 0.5 to comply with Hirschman's entropic uncertainty principle. The quantities $\Delta I_u(f, \Delta f)$, $\Delta I_U(f, \Delta f)$, and $\Delta I_\rho(f, \Delta f)$ when represented in the $(f, \Delta f)$ plane as color-coded map constituting the (SE infogram), (SES infogram), and the (average infogram), respectively [22].

2.4. Dispersion entropy

Sample entropy and permutation entropy are widely used as indicators of time series fluctuation [27-30]. Although PE is computationally fast and simple, it does not consider the differences between amplitudes values, therefore important information could be lost. On the other hand, sample entropy is a computationally demanding algorithm. As an alternative to these two entropies, dispersion entropy, which can tackle the aforementioned limitations, is introduced by Rostaghi and Azami [25]. DE is faster than PE and can track the amplitude variation in the signal. Moreover, DE sensitivity to noise and minor fluctuations is adjustable through the selection of the number of classes and embedding dimension [25]. Dispersion entropy is derived from Shannon entropy and symbolic dynamics or dispersion patterns [31]. The concept of symbolic dynamics is devised from coarse-graining of the time series, i.e., transforming the measurements into a new signal with few different elements by a certain nonlinear mapping process. Hence, the dynamics of a signal can be studied by tracking the distribution of the symbolic sequences or dispersion patterns. Indeed, some minor details may be lost, but some invariant dynamic properties can be kept and detected more easily. In order to understand how DE works and compare it with PE, the steps required to obtain the DE for a time series $x(n)$, will be provided first.

Step 1: Mapping the data to a new time series using normal cumulative distribution function:

$$y(n) = \frac{1}{2}\left(1 - \mathrm{erf}\left(-\frac{x'(n)}{\sqrt{2}}\right)\right) \qquad \mathrm{n} = 0, 1, 2, \tag{7}$$

Where: erf(.) represents error function, $x'(n)$ is the signal $x(n)$ normalized to zero mean and unity variance, and the resulting data $y(n)$ is ranging from 0 to 1.

Step 2: Map $y(n)$ to a new data $z(n)$ ranging from 1 to c: $z(n) = \mathrm{round}\,(c\, y(n) + 0.5)$, where c is the number of classes.

Step 3: Choosing the embedding dimension m, which represents the number of successive classes taken at a time, and time delay d, the number of possible dispersion patterns is equal to cm, $\pi_i = \{\pi_1, \pi_2, \pi_3, \ldots\}$ where $i = 1, 2, 3 \ldots cm$. Dispersion patterns correspond to vectors whose values are combinations of classes taken m at a time. Many researchers [6, 25-32, 33] recommended using the number of classes c between 4 and 8, delay $d = 1$ and m between 2 and 5.

Step 4: Using a step of d, scan $z(n)$ progressively to find vectors that match each dispersion pattern and count the number of occurrences for each pattern. The number of vectors that can be extracted from the signal is equal to $L - (m - 1)\, d$. Hence, find the probability of each pattern πi:

$$p(\pi_i) = \frac{\text{Number of occurance of } \pi_i}{L - (m - 1)\, d} \tag{8}$$

Where: $p(\pi_i)$ is the probability of each pattern π_i, L is number of samples, m is embedding dimension, and d is delay time.

Step 5: Calculate the dispersion entropy DE (not required for RDE):

$$\mathrm{DE} = -\sum_{i=1}^{c^m} p(\pi_i) \log\left(p(\pi_i)\right) \tag{9}$$

When all the elements of a signal are mapped into one class, the signal is perfectly predictable and, hence, has minimum entropy that equals zero. On the opposite side, if all possible dispersion patterns exist and have the same probability, then, the signal is unpredictable and has maximum entropy value that equals $ln(cm)$. As a numerical example, suppose we have three classes $c = 3$, embedding dimension $m = 2$ and delay $d = 1$. The number of possible dispersion patterns is 32 = 9, $\pi_i = \{\pi_1, \ldots \pi_9\}$ with π_1 {1, 1}, π_2 {1, 2}, π_3 {1, 3}, π_4 {2, 1}, π_5 {2, 2}, π_6 {2, 3}, π_7 {3, 1}, π_8 {3, 2}, π_9 {3, 3}.

Let us assume a time series which is mapped into $z(n) = [3, 1, 2, 3, 3, 2, 1, 1, 2, 2, 1, 3, 3, 3]$. The number of samples is $L = 14$ and the number of vectors that can be extracted is $L - (m - 1)\, d = 13$. When $z(n)$ is scanned for patterns, the number of occurrences of the pattern π_1 is found to be 1, $p(\pi_1) = 1/13$. Similarly, other probabilities are found to be $p(\pi_2) = 2/13$, $p(\pi_3) = 1/13$, $p(\pi_4) = 2/13$, $p(\pi_5) = 1/13$, $p(\pi_6) = 1/13$, $p(\pi_7) = 1/13$, $p(\pi_8) = 1/13$, and $p(\pi_9) = 3/13$.

Hence, $\mathrm{DE} = -\sum_i p(\pi_\mathrm{i})\, ln\, p(\pi_\mathrm{i}) = 2.0981$

It is clear that DE tracks the transitions that occurred in the values of the sequence and it distinguishes between different values. On the contrary, PE considers the ordinal structures of the signal irrespective of what the values are. For example, the sequences [34, 35] and [36, 37] have the same permutation pattern "01" which indicates that the first value is less than the second one. Furthermore, if there exist two or more elements that have the same value in a sequence, the pattern will be determined by their index in ascending order. For these reasons, some important information may be lost.

2.5. Reverse dispersion entropy

Dispersion entropy is to Shannon entropy and cannot be used directly to measure the non-stationarity of the signal. Hence, reverse DE is used, which is analogous to negentropy. Instead of negating Shannon entropy to obtain an indicator that is directly proportional to information content, the distance from the white noise is used. RDE combines the advantages of both DE and reverse permutation entropy RPE [38]. It is defined as the distance of signal probabilities from the white noise probability. The white noise has equal probability $p(\pi_i)$ for all patterns. Since there are cm patterns and the total number of vectors is $L - (m - 1)$ d, the probability of any pattern for a white noise signal is given by:

$$\text{Pwns}\ (\pi_i) = \frac{\dfrac{L-(m-1)d}{C^m}}{L-(m-1)d} = \frac{1}{C^m} \tag{10}$$

Where: statistically, the RDE can be recognized as the square deviation of the probability function $p(\pi_i)$ from the white noise probability, which can be expressed as follows:

$$\text{RDE} = \sum_{i=1}^{C^m}\left(P(\pi_i) - \frac{1}{C^m}\right)^2 = \sum_{i=1}^{C^m} P(\pi_i)^2 - \frac{1}{C^m} \tag{11}$$

For a white noise signal, the RDE is minimum and approaches 0. Whereas, when the signal is constant and, hence, there is only one pattern, the RDE approaches $(1 - 1/cm)$. Therefore, the normalized RDE, which accounts for the number of patterns cm, can be calculated as follows:

$$\text{RDE} = \frac{1}{1-\frac{1}{C^m}} \sum_{i=1}^{C^m} P(\pi_i)^2 - \frac{1}{C^m} \tag{12}$$

The normalized RDE value ranges from 0 for pure white noise to 1 for a constant value signal. Any signal with predictable behavior will exhibit a high RDE, whereas a random unpredictable signal will show a small value of RDE. The normalized RDE will be used throughout the paper. It is worth mentioning that Step 5 of DE calculation is not needed to calculate the RDE. The complete flowchart of RDE including DE is shown in Fig. 1.

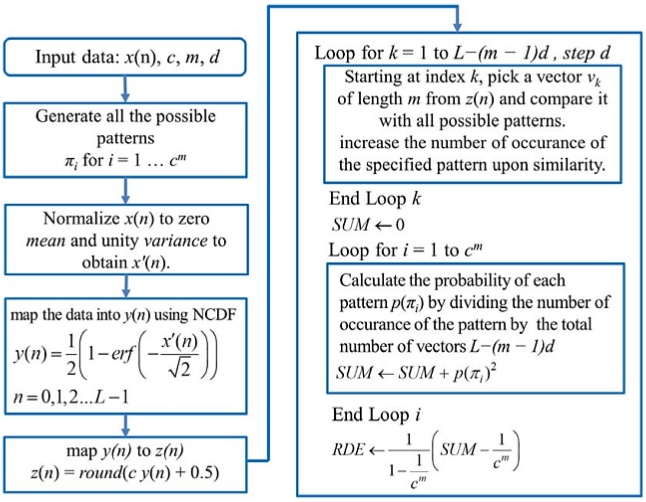


Fig. 1 the flowchart of reverse dispersion entropy (RDE) including dispersion entropy (DE).

2.6. Application to vibration signal

The SE calculated from eq. (3) can be interpreted as instantaneous energy flow of the signal filtered with a filter of center frequency f and bandwidth Δf. Like spectral negentropy which is used by Antoni [22], the RDE can be used to obtain the time domain RDEgram which quantifies the dynamics of the SE and it will be denoted by RDEu $(f, \Delta f)$. Likewise, the RDE can be applied to the SES to obtain the frequency domain RDE gram, denoted by RDEU $(f, \Delta f)$. Shortly, it will be shown that both the time domain and frequency domain spectral RDE grams are required to understand signal behavior, hence a weighted average RDE can be used:

$$\text{RDE}_\rho(f, \Delta f) = \rho \text{RDE}_u(f, \Delta f) + (1-\rho)\,\text{RDE}_U(f, \Delta f) \tag{13}$$

Table 1 compares between negentropy and RDE expressions for some signals. In this table, signals are sampled at F_s, $\delta\Delta f$ stands for a temporal impulse of width $1/\Delta f$ and $\Pi\Delta f$ stands for a rectangle of width Δf and unity amplitude. The same assumptions of Antoni [27] regarding the shape of the spectrum for different signals are adapted here for comparison. $r = F_s/\Delta f$, K the number of pulses for Dirac comb, N is the period between each two successive pulses, $L = KN$. The details of these expressions are provided in the proof of eq. (12). The parameters of RDE method are $c = 4$, $m = 2$, and $d = 1$. For Dirac comb function, the values of RDE in the time and frequency domain can be calculated from the following equations:

$$\sigma(L, K, r) = \frac{1}{1-\frac{1}{C^m}}\left[\left(\frac{L-1-2K-K(r-1)^2}{L-1}\right) + 2\left(\frac{K}{L-1}\right)^2 + \left(\frac{K(r-1)}{L-1}\right)^2 - \frac{1}{C^m}\right] \tag{14}$$

Where: L is the length of the points, and, $K = L/N$ is the number of pulses, and r is the pulse width.

Equation (14) is proved in the proof of eq. (12), whereas some of the values are calculated as in Table 1. It can be seen from that for White Complex Gaussian Noise (WCGN) the RDE is 0 while it is 0.4228 for negentropy. This suggests that RDE can handle noise better than negentropy. For a sinusoidal SE (cyclostationary vibration signal), the time domain RDE is generally small and depends on the frequency of the sinusoid and ranging from 0.23 for very low frequency (1 cycle in the entire length) to 0.1 for high-frequency SE. However, using time domain RDE, just like time domain negentropy, is not a good idea for this type of signal. On the other hand, the frequency domain RDE approaches its maximum value for such signals due to the fact that there is only one impulse in the SES.

For Dirac comb (pulse train) SE with pulses spaced at N, the SES is also pulse train spaced at $K = L/N$ according to Fourier transform properties. It can be easily shown that, for certain temporal spacing N, the RDE does not depend on the total length L. In fact, the time domain RDE, like the time domain negentropy, is a function of N only (when considering certain r). However, the RDE is less sensitive to spacing changes. For example, when $L = 4000$ and $r = 1$, if N is changed from 10 to 1000, the negentropy change will be 300% while, the RDE change is only 156%. This implies that RDE is less susceptible to the rate of repetition of transient events. Figure 2 shows the time domain Dirac comb RDE as a function of spacing N for total length $L = 4000$.

Table 1 Negentropies and RED expressions for some singles.

Squared envelope	Squared envelope spectrum	$\Delta I_u(f, \Delta f)$	$\Delta I_U(f, \Delta f)$	$\Delta I_{1/2}(f, \Delta f)$	$\mathrm{RDE}_u(f, \Delta f)$	$\mathrm{RDE}_U(f, \Delta f)$	$\mathrm{RDE}_{1/2}(f, \Delta f)$
WCGN	WCGN	≈ 0.4228	≈ 0.4228	≈ 0.4228	0	0	0
Sinusoidal $A\cos\left(2\pi n\frac{\alpha_0}{F_s}+\phi\right)$	$\frac{L}{2}Ae^{j\phi}\delta(\alpha-\alpha_0)$	0.3068	$\ln(L)$	$\frac{1}{2}\ln(L)$	0.1 to 0.23	≈ 1	0.55 to 0.62
Single impulse $A\delta_{\Delta f}(n-n_0)$	$Ae^{-j2\pi n_0\frac{\alpha}{F_s}}\prod_{\Delta f}(\alpha)$	$\ln\left(\frac{L}{r}\right)$	$\ln(r)$	$\frac{1}{2}\ln(L)$	$\sigma(L, 1, r)$	$\sigma\left(L, \frac{L}{r}, 1\right)\ r>1$	$\frac{1}{2}(\sigma_u+\sigma_U)$
Dirac comb $A\sum_{m=0}^{K-1}\delta_{\Delta f}(n-mN)$	Dirac comb $KA\sum_{p=0}^{N-1}\delta(\alpha-pKF_s)\,.\prod_{\Delta f}(\alpha)$	$\ln\left(\frac{N}{r}\right)$	$\ln(K\,r)$	$\frac{1}{2}\ln(L)$	$\sigma(L, K, r)$	$\sigma\left(L, \frac{N}{r}, 1\right)$	$\frac{1}{2}(\sigma_u+\sigma_U)$

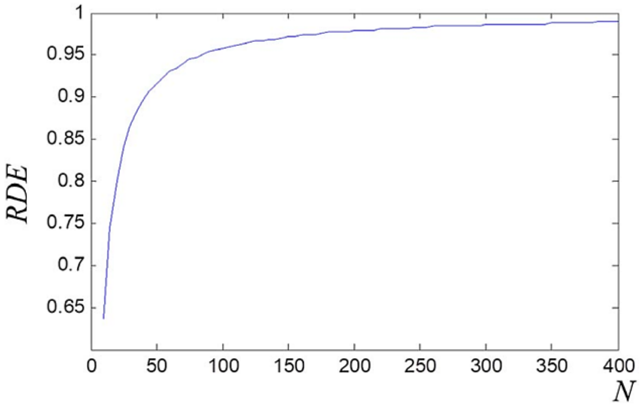


Fig. 2 RED of Dirca comb as function of comb spacing.

2.7. Implementation

Step 1: The SE of the signal, $u(n;\ f, \Delta f)$, is extracted using eq. (3). To ensure proper energy flow sampling and prevent information loss, the value of overlapping parameter R is chosen as $R = Nw/4$ except at the zero frequency and Nyquist frequency where it is taken as $Nw/2$ [22]. Normally the Fast Fourier Transform (FFT) or equivalent algorithm is used to calculate SE. The window length is changed over a number of selected values to provide variable bandwidth Δf.

Step 2: The SE signal $u(n;\ f, \Delta f)$ at a given frequency and frequency resolution is processed using RDE algorithm to find the time domain RDE $u(f, \Delta f)$, which can be represented by a color-coded map RDEgram.

Step 3: The SE signal is DFT-transformed to obtain squared energy spectrum $U(\alpha;\ f, \Delta f)$ which is processed by RDE to obtain frequency domain $\mathrm{RDE}_U(f, \Delta f)$. Efficient FFT algorithm such as FFTW can be used for this purpose. The RDEgram for $\mathrm{RDE}_U(f, \Delta f)$ can be constructed.

Step 4: The combined RDEgram can be evaluated according to eq. (13).

Step 5: FIR filters are designed using the data obtained from RDEgrams. The FIR filter has center frequency corresponding to the regions of highest RDEgram value and filter bandwidth is set to twice the corresponding frequency resolution; Bandwidth = $2\Delta f$.

Step 6: Vibration signal is filtered using the filter designed in Step5 to reveal the transients in the signal. The spectrum of the filtered signal can also be obtained after rectifying or squaring the filtered signal to obtain the envelope. The flowchart of the proposed method is shown in Fig. 3.

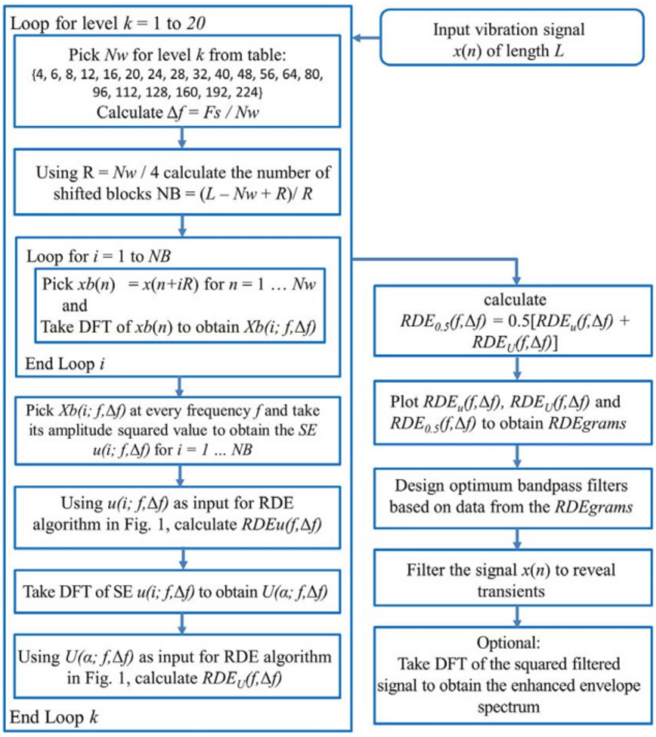


Fig. 3 Flowchart of the used method.

3. Experimental work

3.1. Experimental rig

In this work, the experimental rig consists of electrical motor, gear box, sensors, bearing loader, bearing, data acquisition and computer as shown in Fig. 4.

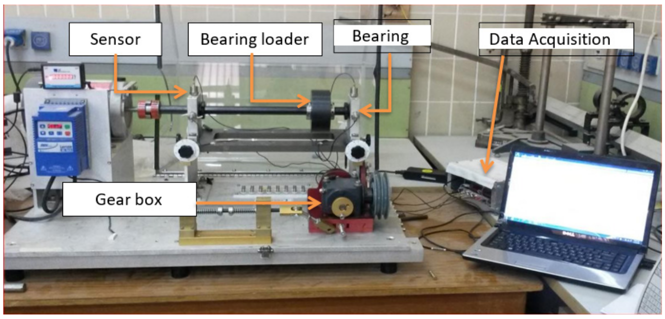


Fig. 4 Experimental rig used in this work.

3.2. Ball bearings

In this work, ball bearings type (KOYO 1205 C3) was used and the specifications of this ball bearing, the ball bearing consists of three parts as shown in Fig. 5.

Fig. 5 Ball bearings type (KOYO 1205 C3).

3.3. Experimental procedure

The experimental procedure of this work is:

1. Select the ball bearing specimen test (intact, pit, slot and scratch) and install the ball bearing in the experimental rig.
2. Select the rotation velocity of the electrical motor.
3. The experimental acceleration wave from the vibration sensor (accelorameter) for four to six time steps and each time step = 100 sec, and save the wave in computer.
4. Use the signal tool program to convert the wave from time domain into frequency domain.
5. Convert the saved data into color photos using reversed dispersion entropy (RDE) that is programing using Matlab program.
6. Save the color photos for each type of defects and then use it in the CNN program. In CNN program, 70% of the total color photos are used in training process of CNN, while, the 30% of the total color photos are used in testing process of CNN or checking the CNN results. The specifications of PC computer using in conversion process and the time spend to convert 100 sec of data is about 63 min with overlap about 90%. The time consume for conversion process reduces a new computer with following specifications is used:

4. Results and discussion

In this work, the experimental and theoretical results explained and discussed. This section is divided into two parts according to the rotational velocity used in this work 30 and 25 Hz.

Table 2 lists the max. RDE gram (Normal) values and the corresponding spectral frequency and bandwidth for normal bearings at 30 and 25 Hz, respectively. It is found that the maximum RDE occurs at 1200 Hz spectral frequency. In addition, a general except for readings with frequency 300 Hz is noticed. Figures 4 and 5 show the RDE gram for normal ball bearing at 30 and 25 Hz.

Table 3 lists the max. RDE gram (inner) values and the corresponding spectral frequency and bandwidth for normal bearing at 30 and 25 Hz for bearing kay 1296 *C*3. It is found that the maximum RDE occurs at 1200 Hz and 14000 Hz spectral frequency and a general except for reading with frequency 300 Hz and 2000 Hz when the rotational speed is 30 and 25 Hz, respectively. Figures 6 and 7 show the RDE gram for bearing with inner pit fault at 30 and 25 Hz.

For the outer race pit case, the maximum RDE occurs at 1200 and 16000 Hz spectral frequency and a general except for reading with frequency 300 and 3000 Hz at 30 and 25 Hz respectively. Figures 8 and 9 show the RDE gram for bearing with Outer pit fault at 30 and 25 Hz.

For the ball pit case, Table 4 shows in detail the max. RDE gram (ball) values and the corresponding spectral frequency and bandwidth for normal bearing at 30 and 25 Hz for bearing kay 1296 C3. It is found that the maximum RDE occurs at 1200 and 17000 Hz spectral frequency and a general except for reading with frequency 300 and 4500 Hz, respectively.

For CNN, Fig. 10 shows the training progress for pitting fault at 30 Hz. It is clear that after 80 iterations, the training accuracy starts to stabilize and the loss becomes very small. From confusion matrix in Fig. 11, it is clear that the CNN achieved accuracy of 100%. In the other word, Fig. 12 shows the training progress for pitting fault at 25 Hz. It is clear that after 80 iterations, the training accuracy starts to stabilize and the loss become very small. From confusion matrix in Fig. 13, it is clear that the CNN achieved accuracy of 100%.

Table 2 the max. RDE gram (Normal) values and the corresponding spectral frequency and bandwidth for normal bearing at 30 and 25 Hz for bearing koyo 1205 C3.

30 Hz				25 Hz			
Block Number	Max RDE	Spectral Freq.	Bandwidth	Block Number	Max RDE	Spectral Freq.	Bandwidth
1	0.15173	1000	1000	1	0.16976	1200	1200
2	0.15968	1200	1200	2	0.16342	1200	1200
3	0.15514	1000	1000	3	0.1627	1200	1200
4	0.15824	1200	1200	4	0.15859	1200	1200
5	0.16316	1200	1200	5	0.15466	1200	1200
6	0.16492	1200	1200	6	0.16167	1200	1200
7	0.1633	1200	1200	7	0.16065	1200	1200
8	0.16433	1200	1200	8	0.1582	1200	1200
9	0.17222	1200	1200	9	0.15597	1200	1200
10	0.16863	1200	1200	10	0.15705	1200	1200
11	0.16817	3000	428.571	11	0.16374	1200	1200
12	0.16415	3000	428.571	12	0.16207	1200	1200
13	0.1701	3000	428.571	13	0.16656	1200	1200
14	0.1536	3000	428.571	14	0.16468	1200	1200
15	0.16723	3000	428.571	15	0.16767	1200	1200
16	0.15924	3000	428.571	16	0.16735	1200	1200
17	0.15542	3000	428.571	17	0.16657	1200	1200
18	0.15072	3000	428.571	18	0.16791	1200	1200
19	0.15072	3000	428.571	19	0.16368	1200	1200
20	0.15729	3000	428.571	20	0.17426	1200	1200
21	0.14943	3000	428.571	21	0.15053	300	300
22	0.13996	1200	6000	22	0.12773	300	300
23	0.14517	3000	500	23	0.1456	1200	300
24	0.15161	3000	500	24	0.16032	1200	300
25	0.14853	3000	428.571	25	0.16369	1200	300
26	0.15173	3000	429.571	26	0.17831	1200	300
27	0.15579	3000	430.571	27	0.1925	1200	300
28	0.15589	3000	375	28	0.21685	1200	300
29	0.16302	3000	428.571	29	0.22369	1200	300
30	0.16696	3000	429.571	30	0.15053	300	300
31	0.14943	3000	428.571	31	0.14074	1200	1200
32	0.13996	12000	6000	32	0.14352	1200	1200
33	0.14517	3000	500	33	0.1487	1200	1200
34	0.15161	3000	500	34	0.14255	1200	1200
35	0.14853	3000	428.571	35	0.14286	1200	1200
36	0.15173	3000	428.571	36	0.14638	1200	1200
37	0.15579	3000	428.571	37	0.14933	1200	1200
38	0.15589	3000	375	38	0.14615	1200	1200
39	0.16302	3000	428.571	39	0.15626	1200	1200
40	0.16696	3000	428.571	40	0.15204	1200	1200

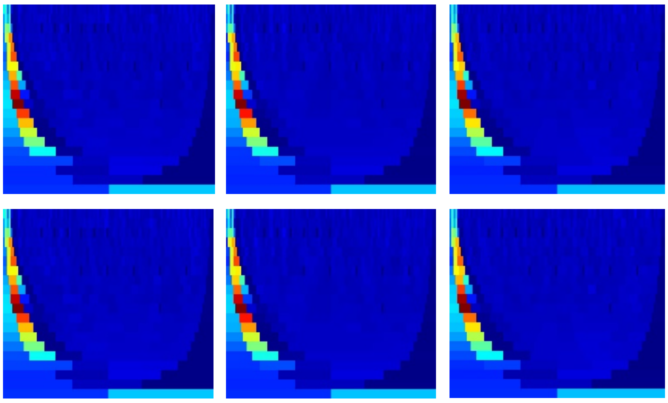


Fig. 6 the RDE gram for bearing with normal pit fault at 30 Hz.

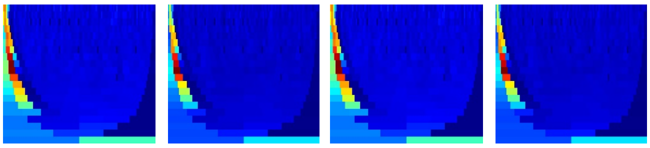


Fig. 7 the RDE gram for bearing with normal pit fault at 25 Hz.

Table 3 the max. RDE gram (Inner) values and the corresponding spectral frequency and bandwidth for normal bearing at 30 and 25 Hz for bearing koyo 1205 C3.

30 Hz				25 Hz			
Block Number	Max RDE	Spectral Freq.	Bandwidth	Block Number	Max RDE	Spectral Freq.	Bandwidth
1	0.10946	3857.142	214.2857	1	0.15472	4500	1500
2	0.11602	3857.142	215.2857	2	0.15134	4500	1500
3	0.1207	5625	375	3	0.15174	4500	1500
4	0.12806	5625	375	4	0.15199	4500	1500
5	0.11429	3857.142	214.2857	5	0.14987	4500	1500
6	0.11115	3857.142	214.2857	6	0.15039	4500	1500
7	0.11111	3857.142	214.2857	7	0.15042	4500	1500
8	0.11239	3857.142	214.2857	8	0.14775	4800	1200
9	0.11662	3857.142	214.2857	9	0.15039	4500	1500
10	0.10998	3857.142	214.2857	10	0.14719	4500	1500
11	0.15896	18000	2000	11	0.14887	1714.285	1714.285
12	0.15994	17142.85	1714.285	12	0.15377	1714.285	1714.285
13	0.16255	16800	1200	13	0.15229	1714.285	1714.285
14	0.16401	16800	1200	14	0.15089	1714.285	1714.285
15	0.16133	16800	1200	15	0.14967	1714.285	1714.285
16	0.16429	16800	1200	16	0.14749	1714.285	1714.285
17	0.15967	6000	6000	17	0.15012	2000	2000
18	0.15752	6000	6000	18	0.15061	1714.285	1714.285
19	0.15864	6000	6000	19	0.14367	1714.285	1714.285
20	0.15991	16800	1200	20	0.14909	1714.285	1714.285
21	0.12419	3857.142	214.2857	21	0.15259	1714.285	428.5714
22	0.12628	3857.142	214.2857	22	0.15407	1714.285	428.5714
23	0.11789	5625	375	23	0.15331	1714.285	428.5714
24	0.12316	5625	375	24	0.15218	1714.285	428.5714
25	0.11566	5625	375	25	0.14557	1714.285	428.5714
26	0.11982	3750	250	26	0.14005	5750	250
27	0.12412	3857.142	214.2857	27	0.1408	5750	250
28	0.11733	3750	250	28	0.14123	5750	250
29	0.12222	3857.142	214.2857	29	0.14135	6000	1500
30	0.12416	3857.142	214.2857	30	0.13634	5750	250
31	0.12403	3857.142	214.2857	31	0.1364	14000	1000
32	0.12599	3857.142	214.2857	32	0.1459	14000	1000
33	0.11742	5625	375	33	0.14209	14000	1000
34	0.12396	5625	375	34	0.14063	13714.28	857.1429
35	0.11987	3857.142	214.2857	35	0.14063	14000	1000
36	0.11805	3750	250	36	0.15393	14000	1000
37	0.1218	3750	250	37	0.15393	13714.28	857.1429
38	0.11827	3857.142	214.2857	38	0.14914	13714.28	857.1429
39	0.12441	3857.142	214.2857	39	0.15845	13714.28	857.1429
40	0.12388	3857.142	214.2857	40	0.1491	4800	1200

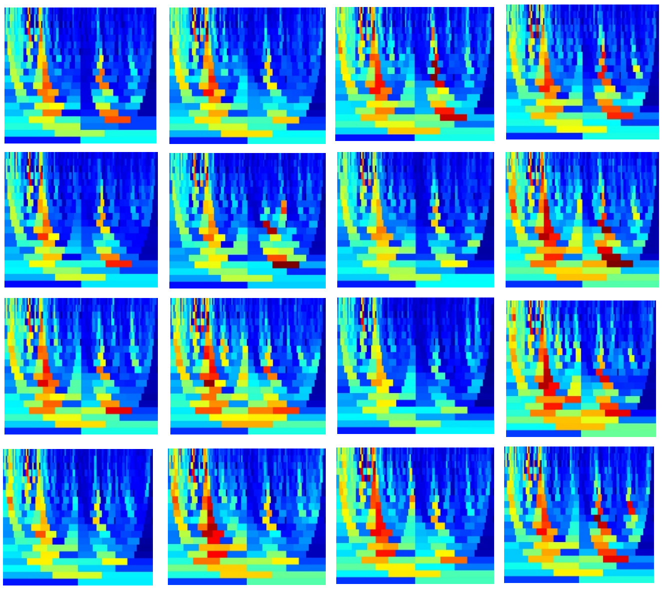


Fig. 8 the RDE gram for bearing with inner race pit fault at 30 Hz.

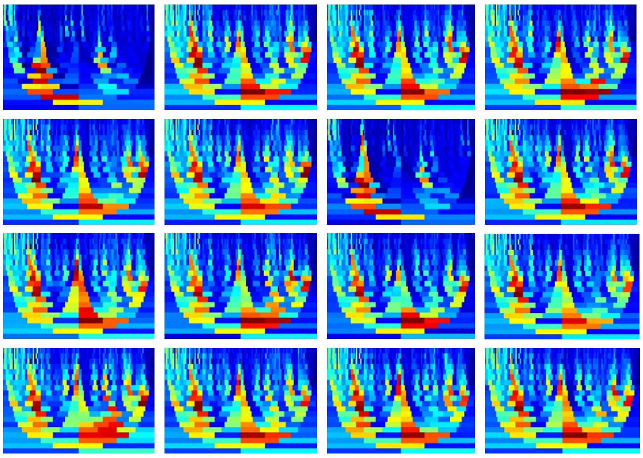


Fig. 9 the RDE gram for bearing with inner race pit fault at 25 Hz.

Table 4 the max. RDE gram (Outer) values and the corresponding spectral frequency and bandwidth for normal bearing at 25 Hz for bearing kay 1296 C3.

30 Hz				25 Hz			
Block Number	Max RDE	Spectral Freq.	Bandwidth	Block Number	Max RDE	Spectral Freq.	Bandwidth
1	0.14712	5142.85	1714.28	1	0.13851	16000	4000
2	0.15041	5142.85	1714.28	2	0.13813	16000	4000
3	0.15119	5142.85	1714.28	3	0.13985	16000	4000
4	0.15202	5142.85	1714.28	4	0.13885	16000	4000
5	0.14871	5142.85	1714.28	5	0.1403	16000	4000
6	0.14914	5142.85	1714.28	7	0.14876	16000	4000
7	0.14918	5142.85	1714.28	8	0.14662	16000	4000
8	0.1458	5142.85	1714.28	9	0.14551	16000	4000
9	0.14866	5142.85	1714.28	10	0.14029	16000	4000
10	0.15195	16000	4000	11	0.15085	15000	3000
11	0.1512	5142.85	1714.28	12	0.15197	16000	4000
12	0.15166	4800	1200	13	0.13688	16000	4000
13	0.15039	4800	1200	14	0.13662	16000	4000
14	0.15225	5142.85	1714.28	15	0.13431	16000	4000
15	0.14914	5142.85	1714.28	7	0.14876	16000	4000
16	0.14993	5142.85	1714.28	16	0.13222	16000	4000
17	0.15162	5142.85	1714.28	17	0.12898	16000	4000
18	0.15152	5142.85	1714.28	18	0.12739	16000	4000
19	0.15326	4500	1500	19	0.1256	16000	4000
20	0.15382	4500	1500	20	0.12576	16000	4000
21	0.15225	5142.85	1714.28	21	0.15113	16000	4000
22	0.14968	4800	1200	22	0.15125	16000	4000
23	0.15277	5142.85	1714.28	23	0.13596	16000	4000
24	0.14893	5142.85	1714.28	24	0.1362	16000	4000
25	0.15198	5142.85	1714.28	25	0.13375	16000	4000
26	0.15213	5142.85	1714.28	26	0.13397	16000	4000
27	0.15045	5142.85	1714.28	27	0.1298	16000	4000
28	0.1507	5142.85	1714.28	28	0.12852	16000	4000
29	0.15306	5142.85	1714.28	29	0.12436	16000	4000
30	0.15135	5142.85	1714.28	30	0.12838	16000	4000
31	0.17452	4000	4000	31	0.14102	16000	4000
32	0.17252	4000	4000	32	0.13972	16000	4000
33	0.16535	4000	4000	33	0.1413	16000	4000
34	0.20912	4000	4000	34	0.14197	16000	4000
35	0.20807	16800	2400	35	0.13839	16000	4000
36	0.20875	16800	2400	36	0.13618	16000	4000
37	0.19356	16800	2400	37	0.13164	16000	4000
38	0.199	16800	2400	38	0.13496	16000	4000
39	0.20732	16800	2400	39	0.13142	16000	4000
40	0.20362	16800	2400	40	0.13049	16000	4000

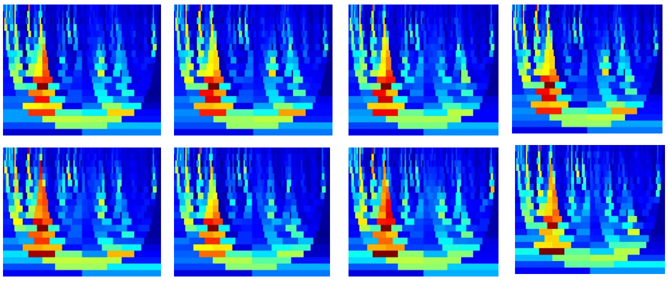


Fig. 10 a sample of the outer pitting RDE gram.

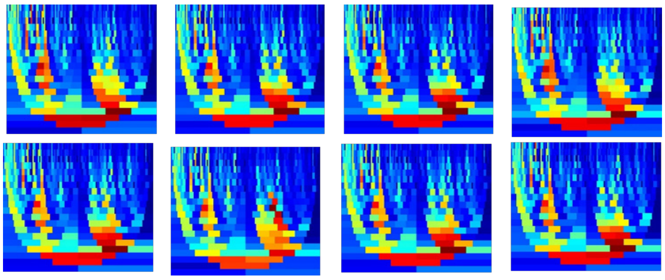
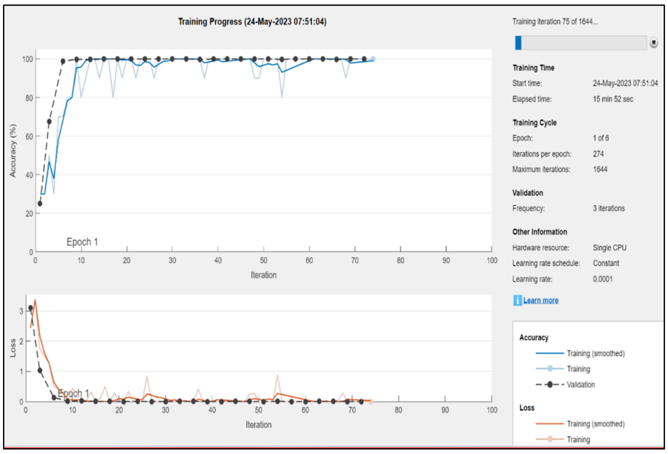


Fig. 11 the RDE gram for bearing with outer race pit fault at 25 Hz.



Training iteration 75 of 1644...

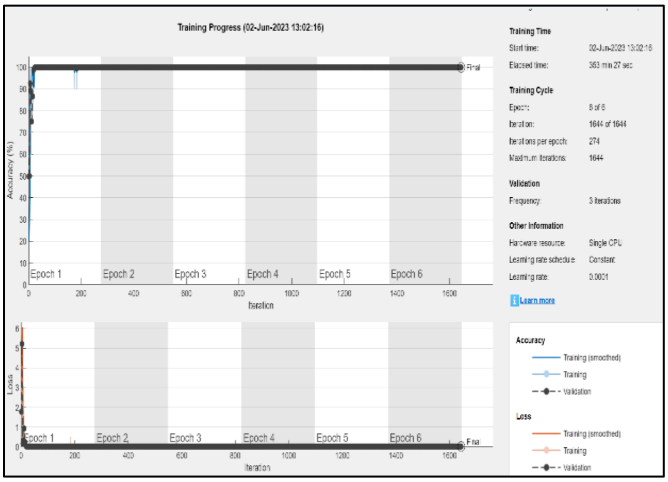
Training Time
Start time: 24-May-2023 07:51:04
Elapsed time: 15 min 52 sec

Training Cycle
Epoch: 1 of 6
Iterations per epoch: 274
Maximum iterations: 1644

Validation
Frequency: 3 iterations

Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001

(a)



Training Time
Start time: 02-Jun-2023 13:02:16
Elapsed time: 363 min 27 sec

Training Cycle
Epoch: 6 of 6
Iteration: 1644 of 1644
Iterations per epoch: 274
Maximum iterations: 1644

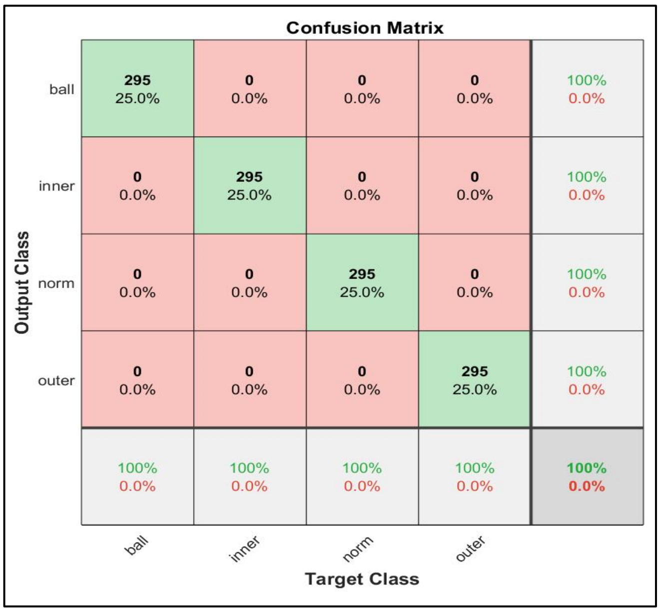
Validation
Frequency: 3 iterations

Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001

(b)

Fig. 12 the training progress for pitting fault at 30 Hz.

Figure 13 shows the confusion matrix at 30 Hz composes four types of input data (normal bearing, defects at ball, defects at inner race, defects at outer race) and four types of output data. The green boxes represent the intersection between the input and output data of the same defect type, while the pink boxes represent the intersection between different defect types. The intersection happened between the same defect types, which is the ball in this case with 25% or 295 images.



Confusion Matrix

Output Class \ Target Class	ball	inner	norm	outer	
ball	295 25.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
inner	0 0.0%	295 25.0%	0 0.0%	0 0.0%	100% 0.0%
norm	0 0.0%	0 0.0%	295 25.0%	0 0.0%	100% 0.0%
outer	0 0.0%	0 0.0%	0 0.0%	295 25.0%	100% 0.0%
	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%

Fig. 13 Confusion matrix at 30 Hz.

Training iteration 51 of 1680...

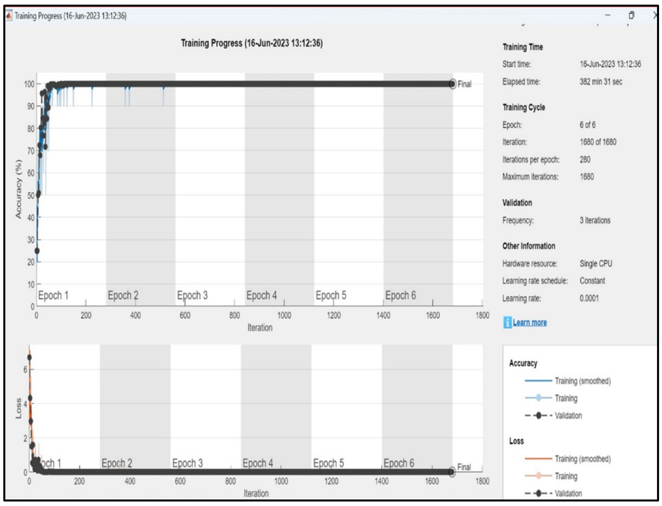
Training Time
Start time: 16-Jun-2023 13:12:36
Elapsed time: 13 min 37 sec

Training Cycle
Epoch: 1 of 6
Iterations per epoch: 280
Maximum iterations: 1680

Validation
Frequency: 3 iterations

Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001

(a)



Training Time
Start time: 16-Jun-2023 13:12:36
Elapsed time: 382 min 31 sec

Training Cycle
Epoch: 6 of 6
Iteration: 1680 of 1680
Iterations per epoch: 280
Maximum iterations: 1680

Validation
Frequency: 3 iterations

Other Information
Hardware resource: Single CPU
Learning rate schedule: Constant
Learning rate: 0.0001

(b)

Fig. 14 the training progress for pitting fault at 25 Hz.

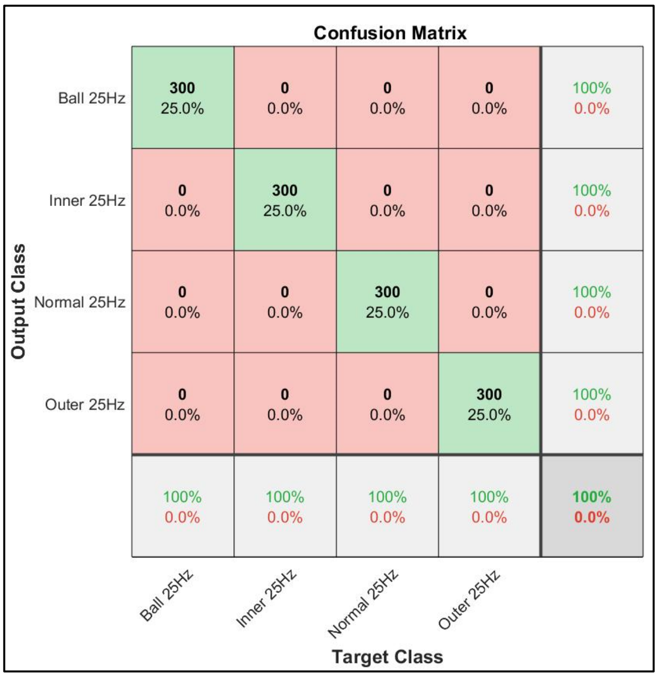


Fig. 15 Confusion matrix at 25 Hz.

7. Conclusions

Based on the current work, it could be concluded that the using a combination RDE-gram and CNN is an efficient technique to predict the catastrophic failure produced in ball bearings before occurring. The accuracy of prediction approaches a level of 100 % for predicting the pit faults at different locations such as inner ring, outer ring, and the ball element. Accordingly, this is of high importance in the industry to save human life and avoid the high cost ensuing from replacing the machines which can be destroyed as a result of accumulative faults of bearing. For these reasons, the demands on such tools have been increased highly in the nowadays.

References

[1] J. Antoni, and R. B. Randall, “Differential diagnosis of gear and bearing faults,” Journal of Vibration and Acoustics, Vol. 124, Issue 2, pp. 165-171, 2002. https://doi.org/10.1115/1.1456906

[2] J. Antoni, and R. B. Randall, “A stochastic model for simulation and diagnostics of rolling element bearings with localized faults,” Journal of Vibration and Acoustics, Vol. 125, Issue 3, pp. 282-289, 2003. https://doi.org/10.1115/1.1569940

[3] J. Antoni, “Cyclostationarity by examples,” Mechanical Systems and Signal Processing, Vol. 23, Issue 4, pp. 987-1036, 2009. https://doi.org/10.1016/j.ymssp.2008.10.010

[4] K. Wang, “Incipient fault detection of gearbox bearings through combined vibration analysis,” Fault Detection, Supervision and Safety of Technical Processes 2006, Vol. 2, pp. 1354-1359, 2007. https://doi.org/10.1016/B978-008044485-7/50228-1

[5] J. Antoni, “Cyclic spectral analysis in practice,” Mechanical Systems and Signal Processing, Vol. 21, Issue 2, pp. 597-630, 2007. https://doi.org/10.1016/j.ymssp.2006.08.007

[6] J. Sun, C. Yan and J. Wen, “Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning,” IEEE Transactions on Instrumentation and Measurement, Vol. 67, Issue 1, pp. 185-195, 2018. https://doi.org/10.1109/TIM.2017.2759418

[7] H. Shao, H. Jiang, X. Li, and S. Wu, “Intelligent fault diagnosis of rolling bearing using deep wavelet auto-encoder with extreme learning machine,” Knowledge-Based Systems, Vol. 140, pp. 1-14, 2018. https://doi.org/10.1016/j.knosys.2017.10.024

[8] W. Zhang, G. Peng, C. Li, Y. Chen and Z. Zhang, “A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals,” Sensors, Vol. 17, Issue 2, 2017. https://doi.org/10.3390/s17020425

[9] W. Zhang, G. Peng and C. Li, “Bearings fault diagnosis based on convolutional neural networks with 2D representation of vibration signals as input,” MATEC Web of Conferences, Vol. 95, p 13001, 2017. https://doi.org/10.1051/matecconf/20179513001

[10] W. Zhang, C. Li, G. Peng, Y. Chen and Z. Zhang, “A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load Mechanical Systems and Signal Processing, Vol. 100, pp. 439-453, 2018. https://doi.org/10.1016/j.ymssp.2017.06.022

[11] R. Huang, Y. Liao, S. Zhang and W. Li, “Deep decoupling convolutional neural network for intelligent compound fault diagnosis,” IEEE Access, Vol. 7, pp. 1848-1858, 2019. https://doi.org/10.1109/ACCESS.2018.2886343

[12] H. Wang, S. Li, L. Song and L. Cui, “A novel convolutional neural network based fault recognition method via image fusion of multi-vibration-signals,” Computers in Industry, Vol. 105, pp. 182-190, 2019. https://doi.org/10.1016/j.compind.2018.12.013

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

[14] S. S. Udmale, S. S. Patil, V. M. Phalle and S. K. Singh, “A bearing vibration data analysis based on spectral kurtosis and ConvNet,” Soft Computing, Vol. 23, pp. 9341-9359, 2019. https://doi.org/10.1007/s00500-018-3644-5

[15] S. S. Udmale, S. K. Singh and S. G. Bhirud, “A bearing data analysis based on kurtogram and deep learning sequence models,” Measurement, Vol. 145, pp. 665-677, 2019. https://doi.org/10.1016/j.measurement.2019.05.039

[16] S. Jiang, J. Xuan, J. Duan, J. Lin, H. Tao, Q. Xia, R. Jing, S. Xiong and T. Shi, “Dual attention dense convolutional network for intelligent fault diagnosis of spindle-rolling bearings,” Journal of Vibration and Control, Vol. 27, Issue 21-22, pp. 2403-2419, 2021. https://doi.org/10.1177/1077546320961918

[17] M. Azamfar, J. Singh, X. Li and J. Lee, “Cross-domain gearbox diagnostics under variable working conditions with deep convolutional transfer learning,” Journal of Vibration and Control, Vol. 27, Issue 7-8, pp. 854-864, 2021. https://doi.org/10.1177/1077546320933793

[18] Y. Cao, Y. Ding, M. Jia and R. Tian, “A novel temporal convolutional network with residual self-attention mechanism for remaining useful life prediction of rolling bearings,” Reliability Engineering & System Safety, Vol. 215, 2021. https://doi.org/10.1016/j.ress.2021.107813

[19] K. F. Al-Raheem, A. Roy, K. P. Ramachandran, D. K. Harrison, and S. Grainger, “Rolling Element Bearing Fault Diagnosis Using Laplace Wavelet Envelope Power Spectrum,” EURASIP Journal on Advances in Signal Processing, Vol. 2007, 073629, 2007. https://doi.org/10.1155/2007/73629

[20] J. Antoni, and R. B. Randall, “The Spectral Kurtosis: Application to the Vibratory Surveillance and Diagnostics of Rotating Machines,” Mechanical Systems and Signal Processing, Vol. 20, Issue 2, pp. 308-331, 2006. https://doi.org/10.1016/j.ymssp.2004.09.002

[21] J. Antoni, “Fast Computation of the Kurtogram for the Detection of Transient Faults,” Mechanical Systems and Signal Processing, Vol. 21, Issue 1, pp. 108-124, 2007. https://doi.org/10.1016/j.ymssp.2005.12.002

[22] J. Antoni, “The Infogram: Entropic Evidence of the Signature of Repetitive Transients,” Mechanical Systems and Signal Processing, Vol. 74, pp. 73-94, 2016. https://doi.org/10.1016/j.ymssp.2015.04.034

[23] M. Rostaghi, and H. Azami, “Dispersion Entropy: A Measure for Time-Series Analysis,” IEEE Signal Processing Letters, Vol. 23, Issue 5, pp. 610-614, 2016. https://doi.org/10.1109/LSP.2016.2542881

[24] C. Bandt, and B. Pompe, “Permutation Entropy: A Natural Complexity Measure for Time Series,” Physical Review Letters, Vol. 88, Issue 17, p. 174102, 2002. https://doi.org/10.1103/PhysRevLett.88.174102

[25] Y. Li, X. Chen, J. Yu, X. Yang, and H. Yang, “The Data-Driven Optimization Method and its Application in Feature Extraction of Ship-Radiated Noise with Sample Entropy,” Energies, Vol. 12, Issue 3, p. 359, 2019. https://doi.org/10.3390/en12030359

[26] A. Delgado-Bonal, and A. Marshak, “Approximate Entropy and Sample Entropy: A Comprehensive Tutorial,” Entropy, Vol. 21, Issue 6, p. 541, 2019. https://doi.org/10.3390/e21060541

[27] M. Cocconcelli, R. Rubini, R. Zimroz and W. Bartelmus, “Diagnostics of ball bearings in varying-speed motors by means of artificial neural networks,” 8th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2011, CM 2011/MFPT 2011, Vol. 2, pp. 760-771, 2011. https://hdl.handle.net/11380/664247

[28] B. Eftekharnejad, M. Alssayh, A. Addali and D. Mba, “Spectral kurtosis applied to acoustic emission in bearings”, 25th International Congress on Condition Monitoring and Diagnostic Engineering, Journal of Physics: Conference Series, Vol. 364, 012038, 2012. https://doi.org/10.1088/1742-6596/364/1/012038

[29] F. Tafinine and K. Mokrani, “Real time automatic detection of bearing fault in induction machine using kurtogram analysis”, The Journal of the Acoustical Society of America, Vol. 132, Issue 5, 2012. https://doi.org/10.1121/1.4758764

[30] B. Q. Chen, Z. S. Zhang, Y. Y. Zi, Z. J. He and C. Sun, “Detecting of transient vibration signatures using an improved fast spatial–spectral ensemble kurtosis kurtogram and its applications to mechanical signature analysis of short duration data from rotating machinery”, Mechanical Systems and Signal Processing, Vol. 40, Issue 1, pp. 1-37, 2013. https://doi.org/10.1016/j.ymssp.2013.03.021

[31] S. Pincus, “Approximate Entropy (ApEn) as a Complexity Measure,” Chaos: An Interdisciplinary Journal of Nonlinear Science, Vol. 5, Issue 1, pp. 110-117, 1995. https://doi.org/10.1063/1.166092

[32] A. Hajnayeb, S. E. Khadem and M. H. Moradi, “Design and implementation of an automatic condition monitoring expert system for ball bearing fault detection,” Industrial Lubrication and Tribology, Vol. 60, Issue 2, pp. 93-100, 2008. https://doi.org/10.1108/00368790810858395

[33] L. Cao, Y. Zhao, G. Yu, S. Chen and X. Su, “Research on Gearbox Fault Detection and Diagnosis Based on Improved Spectral Kurtosis Algorithm,” International Conference on Mechatronics, Control and Automation Engineering, pp. 174-177, 2016. https://doi.org/10.2991/mcae-16.2016.41

[34] T. Mushiri and C. Mbohwa, “Vibration based condition Monitoring of Rotating machinery using fuzzy logic,” Proceedings of the World Congress on Engineering and Computer Science 2016, Vol. 2, pp. 19-21, 2016.

[35] L. G. Deshpand, “Simulation of Vibrations Caused by Faults in Bearings and Gears,” Ph.D. Thesis, University of New South Wales, July, 2014. https://doi.org/10.26190/unsworks/16918

[36] L. M. Hassan, “modeling and experimental investigation of rotating Shafts vibration problems,” M.Sc. Thesis, University of Basrah, 2023.

[37] S. N. Abdulkareem, “Advanced vibration analysis in bearing and gearbox faults detection,” M.Sc. Thesis, University of Basrah, April.

[38] J. S. Richman, and J. R. Moorman, “Physiological Time-Series Analysis using Approximate Entropy and Sample Entropy,” American Journal of Physiology-Heart and Circulatory Physiology, Vol. 278, Issue 6, pp. 2039-2049, 2000. https://doi.org/10.1152/ajpheart.2000.278.6.H2039