Journal of Computers Vol. 35 No. 5, October 2024, pp. 47-60
doi: 10.53106/199115992024103505004

Diagnosis Bearing Faults Based on the Triplet Optimized Embedding Models

Xinyang Chen*

Liuzhou Railway Vocational Technical College, China
1627398192@qq.com

Received 22 September 2023; Revised 26 September 2023; Accepted 27 September 2023

Abstract. In rotating machinery, the bearing is one of the important components which improves the rotating machinery's performance. The bearing quality determines the machine's performance and reliability. Therefore, fault detection is a key technology to ensure the bearing's safety and reliability. In the bearing fault diagnosis, separating the sensitive signal from vibration data is one of the challenging tasks due to the large volume of the rolling bearings. The research challenges are overcome using the Triplet Optimized Embedding Model (TOEM) that classifies the faults bearings with maximum accuracy. The triple embeddings are initially created using the ant-optimized long short-term neural model that minimizes the vibration signal. This process extracts the features from the collected data and has been classified using the autoencoder neural model. Encoder, decoder, and activation functions are incorporated during the classification process to classify the faults in bearings. The training process maximizes the fault detection accuracy compared to the existing machine learning classifiers.

Keywords: rotating machinery, bearings, fault diagnosis, triplet optimized embeddings, ant optimized long short-term neural model, autoencoder, classification

1 Introduction

In rotating machinery [1], the bearing is an important mechanical component that works under various conditions, such as high-impact and high-load rolling bearings. The difficult work conditions cause machine and bearing failures [2], in which 40% of motor failures occurred. The bearing and other faults must be addressed immediately to improve machine performance. The machine requires an effective fault diagnosis system [3] to ensure reliability, accuracy, and robustness. The bearing faults have occurred in any part of the machine, like the cage, ball, and inner and outer races [4]. Currently, several machine learning and learning techniques are incorporated to recognize the faults in earlier stages. Different sensors like stator current, vibrations, acoustic noise, thermal imaging, and multiple sensor fusion observe the machine's signals [5-6]. These methods observe and explore the vibration signal to extract the prominent features. The features are used to understand the vibration signal characteristics and improve the overall fault diagnosis rate. The fault identification process consists of two important steps such as feature derivation [7-8] and fault classification. The feature extraction process derives the vibration signal information like amplitude, frequency information, time, and frequency domain statistical features. These features are extracted with the help of various methods [9-11], such as Hilbert Transform (HT), Fast Fourier Transform (FFT), Short-Time Fourier Transform (STFT), and wavelet transforms are utilized to get the signal features. The extracted features are input to the classifiers to predict the faults. The classification process uses the machine learning technique and learning rules to determine the changes in every vibration signal. The classification is performed with the help of different algorithms [12-13], such as Support Vector Machine (SVM), Artificial Neural Networks (ANN), Deep Learning Networks (DLN), and K-Nearest Neighboring (KNN) approaches. These approaches use the learning algorithms and hyperplane to determine the output value. However, the classification process was influenced due to the noisy information and irrelevant details. The noise details create computation complexities while analyzing the large volume of data.

The rotating machine's information varies depending on working conditions that create non-stable issues [14]. The dynamic change in bearing information causes background noise and high dimensionality issues. Existing machine learning techniques address the research difficulties but require frequent learning and training to improve the overall recognition rate [15]. In addition, the machine learning algorithm requires discriminative

* Corresponding Author

attributes to understand the high-dimensional data. The research difficulties are addressed using deep learning because it can learn and process high-dimensional data. The deep learning process has several layers with activation functions that predict the output accurately and efficiently. However, detecting faults is an essential component of the technology to guarantee the bearing's safety and reliability. Due to the huge volume of rolling bearings, one of the challenging tasks in bearing defect diagnostics is isolating the sensitive signal from the vibration data. It is one of the tasks that can be difficult. The research issues are addressed with the help of the Triplet Optimized Embedding Model (TOEM), which classifies the fault bearings with maximum accuracy. The TOEM approach uses the three methodologies in embedding concepts to improve the overall recognition accuracy. The wavelet denoising process initially decomposes the signal to identify the low coefficients. The data segmentation is performed to split the dataset into a subset that simplifies the high-dimensional data computation. Afterward, a long, short-term neural network extracts the features to determine the faults from the vibrating signals. The feature extraction process is optimized using an ant colony approach that fine-tunes the network parameters. Finally, the autoencoder reconstructs the inputs to identify the faults with maximum accuracy and solve the research problems. The overall objective of this study is listed as follows.

- To analyze the vibrating signals using an autoencoder neural model to classify the faults and normal bearing features with maximum diagnosis accuracy,
- To design the wavelet denoising and thresholding procedure to remove the outliers and irrelevant information to maintain reliability while analyzing large volumes of data.
- To develop the ant-optimized long short-term memory network for extracting the vibrating signal features to minimize the deviation between the output values.

Then, the rest of the paper is organized as follows: section 2 analyzes the various researcher's opinions regarding the fault diagnosis process. Section 3 describes the TOEM method's working process, and the system's efficiency is evaluated in Section 4. The conclusion is discussed in section 5.

2 Related Works

Support Vector Machine (SVM) and Convolution Neural Networks (CNN) are suggested to diagnose the bearing faults. The SVM and CNN approach overcame the classification problems while analyzing high-dimensional data [16]. The collected machine signals are processed with the help of the continuous wavelet transform that changes the signals into two-dimensionality frequency images. The converted images are fed into the constructed model, which uses the CNN approach to locate the fault region. The integrated SVM and CNN approach uses the MFTP and CWRU datasets in which the introduced system attains high accuracy. However, the system requires optimization methods to reduce feature redundancy and flexibility issues.

Rolling bearing faults are identified by extracting the multiscale local features using the Back Propagation Neural Networks (BPNN) ([17]. The frequency and time domain information are explored to extract the discriminative features. The extracted features are classified with the help of BPNN, which identifies the faults with maximum accuracy. The analysis uses the three-dimensional level of the discrete wavelet transform to compose the original signal into a sub-signal. The decomposed signals are analyzed to get the multiscale local features. The derived features point out the fault region effectively.

Deep Neural Networks (DNN) are used to detect faults in rolling bearings [18]. The DNN approach uses metric and feature learning to analyze the fault location. Feature learning uses twin neural networks to identify the sample pairs to predict the features. Then, metric learning is applied to determine the sample pair similarities integrated with the test sample to enhance the overall performance. Finally, label smoothening is incorporated to maximize the classification process, and the DNN model predicts the faults in cross-bearing locations and working conditions.

Optimal Ensemble Deep Transfer Networks (OSETN) is applied to identify the rolling bearing faults [19]. The OSETN approach has three learning processes: domain adaptation, ensemble, and transfer learning. The maximum mean discrepancy is initially integrated with the deep transfer networks to determine the feature adaptation. Then, transfer learning is utilized to predict the start point. At last, an ensemble learning process was incorporated to maximize the overall fault detection accuracy. During the analysis, a particle swarm optimization algorithm is applied to update the network parameters that select the network weights depending on voting. The ensemble and optimal strategies reduce fault diagnosis deviations and improve detection accuracy.

Rotating machinery faults were detected using Bayesian Optimised Convolutional Neural Networks (BOCNN) [20]. The introduced approach concentrates on the network hyperparameters to reduce the deviation between the output computation. The machine axes' accelerometer signals are utilized as the input to the network. The network fine-tunes the parameters using the Bayesian optimization because the inputs are 12,800*1*3 vibration signal matrix. After updating the network parameters, the classification is performed for two rotational speeds and eight different machine states.

Simulated Annealing Optimised Convolutional Neural Networks (SAO-CNN) has also been recommended for application [21]. Initially, the bearing vibration signals are collected, which are processed with the help of the wavelet packet transform that produces the spectrogram. The attained output is processed with the help of the CNN network, which identifies the fault locations. During the analysis, SAO is applied to fine-tune the network parameters. The discussed methods use the Case Western Reserve University database information to evaluate the system performance. The effective utilization of the optimization algorithm reduces the overall fault diagnosis deviations and maximizes the recognition accuracy.

Optimized Adaptive Deep Belief Networks (OADBN) is introduced to address the problems while extracting features from the bearing vibration signals [22]. The signals are processed using the DBN approach with stochastic gradient descent approach-based training data. Backpropagation neural networks are applied to update the network hyperparameter, minimizing overall computation difficulties. In addition, the Slap Swarm optimization algorithm is applied to update the network parameter, minimizing deviation errors and improving the overall recognition accuracy.

Lightweight neural networks (LNN) are maintained for the robustness of fault detection [23]. The system intends to reduce the difficulties in high-requirement processing, poor noise resistance, and artificial feature extraction processes. The gathered vibrated signal is processed with the help of stacked inverted residual convolution networks that maximize the fault diagnosis speed and robustness. This process identifies the fault type, location, and severity with maximum recognition accuracy.

Reinforcement Neural Architecture Search (RNAS) is applied to diagnose the rolling bearing faults [24]. The RNAS has two components: child models and controller models. The gathered machinery vibrate signals are processed using recurrent networks that generate the controllers with a series of actions. The actions are further analyzed to generate the child model to respective choice design. After that, the reinforcement learning process updates the controller parameters directly proportional to the child models. The effective generation of controller and child components reduces the difficulties in the fault diagnosis process.

Fault diagnosis accuracy is maximized with the help of the Multiscale and Multi-Sensor Model [25]. Initially, the multi-sensor is utilized to gather the vibration signals, and a correlation kurtosis weighted fusion process was applied to fuse the signal in three dimensions. The twofold multiscale approach is applied to derive the multiscale features, which is done with the help of deep residual conventional networks. The dilated convolution and multiscale fusion process effectively generate shallow and deep features. Thus, the system ensures maximum fault diagnosis accuracy.

Joint-Loss Convolution Neural Networks (JLCNN) is applied to recognize the bearing faults with a maximum prediction rate [26]. The system intends to resolve the feature extraction problems by constructing the CNN according to the joint-loss concept. The joint-loss approach is used to optimize and help to learn the features. The optimization parameters minimize the computation cost and overfitting risk, which is directly proportional to the prediction accuracy. Then, the system ensures 82.7% accuracy while predicting the bearing faults.

A wide convolution kernel with Convolution Neural Networks is recommended to identify the bearing faults with minimum computation complexity [27]. The gathered original vibrate signal is processed with the help of CNN to extract the features. The wide kernels work on the input to derive the features, which are processed using the multi-layer nonlinear mapping process that derives the features with minimum timelines.

Two palmprint coding systems, PalmHash Code and PalmPhasor Code, are offered tbalance confidentiality and validation performance [28]. Cancelable palmprint programming schemes are expanded from a single to a double dimensionality to lessen the demand for processing power and storage capacity. Additionally, two methods are used to boost the efficiency of 2D cancelable palmprint codes: vertical configuration translation and multi-orientation score level fusion. The experimental results and analyses show that the two measures improve the effectiveness and security of the 2D PalmHash Code and the 2D PalmPhasor Code.

The two-dimensional discrete cosine transform (2DDCT) is proposed to extract features from left and right palmprints, creating a dual-source space [29]. The normalization process reduces high absolute terms coefficient noise in dual-source space. Dual-source discrimination power analysis (DPA) retains and recovers more discriminative coefficients, improving accuracy. Researchers used a contactless palmprint database to prove that dual-source DPA, tailored for multi-instance palmprint feature fusion recognition, is better than single-source DPA.

Conjugate 2DPalmHash Code (CTDPHC) is introduced for cancellable multi-modal biometrics that combines palmprint and palmvein 2DPalmHash Codes (2DPHCs) [30]. Different fusion rules at the score level are analyzed and discussed to find the best fusion approach of 2DPHCs of palmprint and palmvein. In addition, the transposition orientation ranges of 2DPHCs are adjusted for optimal performance. CTDPHC has better verification accuracy and anti-counterfeit ability than 2DPHC without increasing computing complexity or storage cost.

According to the various researcher's opinions, methodology design, and frameworks clearly states that neural networks, especially Deep Neural Networks (DNN) and Convolution Neural Networks (CNN), are widely applied to detect rolling bearing faults. In addition, several optimization algorithms, such as grey wolf, genetic, particle swarm optimization, and simulated annealing approaches, are used to fine-tune the network parameters. However, the existing systems are facing difficulties while separating the sensitive signal from vibration data is one of the challenging tasks due to the large volume of the rolling bearings. The research issues are addressed with the help of Triplet Optimized Embedding Models (TOEM) that identify the faults with maximum accuracy and reliability. Then, the detailed working process of the TOEM is discussed in the below section.

3 Rolling Bearing Fault Diagnosis using Triplet Optimized Embedding Models (TOEM)

The main purpose of this study is to identify the faults in the rolling bearing machinery with maximum recognition accuracy. The faults in the balls' inner and outer bearing rings create production issues. Therefore, the fault diagnosis played an important role in identifying the faults in earlier stages for improving the overall machinery performance. The vibration sensors are initially applied to gather the system vibrations, and the acquired signal is analyzed to predict the faults or damages in various bearing parts. The bearing produces a certain level of vibration on every process and rotation. If the machinery is affected by faults, the bearing will produce high-speed rotation because of the machining errors and shaft deflection. In addition, the bearing fault datasets consist of a small dataset, which is difficult to guess the various part's failures.

The research challenges are overcome by applying the Triplet Optimized Embedding Models (TOEM). The machinery vibrates signal is gathered and processed with the help of the ant-optimized long short-term neural model that extracts the features. The derived features as input are fed into the Autoencoder Neural Model (ANM) to perform the classification process. The ANM model uses the activation function and training process to improve the fault classification efficiency. Then, the overall working process of TOEM is illustrated in Fig. 1.

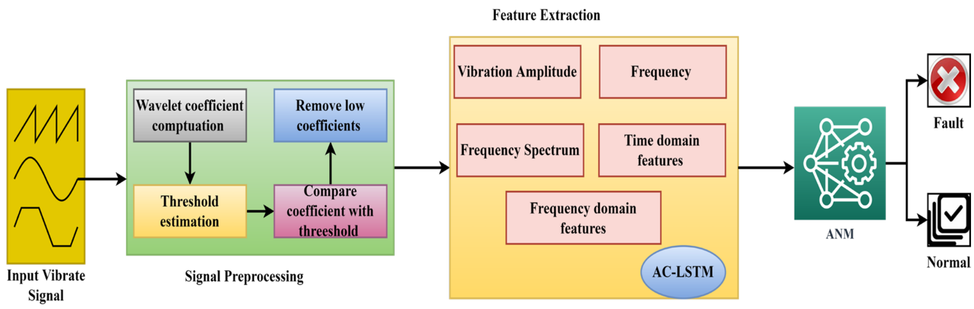


Fig. 1. Process of TOEM-based fault diagnosis system

3.1 Data Collection

This study uses the Case Western Reserve University bearing center [31] information is utilized for research purposes. The dataset consists of faulty and normal bearing information. During the analysis, acceleration data and two hp Reliance electric motor information are examined at a remote location in motor bearings. Electro Discharge Machine (EDM) is utilized in the motor bearing for seeding the faults with a range from 0.007 inches to 0.040 inches in diameter. These faults are included at the rolling element, inner and outer raceway. Therefore, the dataset has vibration data collected from various rolling bearings with various fault types. The data has been gathered from various loads and speeds, including faults and normal bearings. The data is gathered at 12,000 samples/second at normal bearing, 48000 samples/second at single-point end, and 12000 samples/ second at fan end defects. Therefore, the dataset has driver-end accelerometer data (DE), base accelerometer data (BA), fan-end accelerometer data (FE), RPM during testing (RPM), and time series data. Then, the sample dataset information is illustrated in Table 1.

Table 1. Sample dataset information

	Dataset	Diameter (mm)	Depth (mm)	Motor speed (rpm)	Motor load (kW)
Normal	4	-	-	1797	0
				1772	0.746
				1750	1.491
				1730	2.237
Minor faults	Inner race (4) Ball (4) Outer race (4)	0.18	0.28	1797	0
				1772	0.746
				1750	1.491
				1730	2.237
Major fault	Inner race (4) Ball (4) Outer race (4)	0.53	0.28	1797	0
				1772	0.746
				1750	1.491
				1730	2.237

3.2 Data Cleaning and Preprocessing

This work's first step is to remove noise from the collected vibrate data of the rolling machinery. Table 2 vibration data has been gathered from rolling bearings at various operating conditions. Automatic systems use the data to identify faults with maximum recognition accuracy. The vibration data may have several outliers, reducing the overall fault diagnosis process and efficiency. Therefore, the outliers and noise have to be removed from the original data, which is done by applying wavelet denoising. The denoising process enhances the vibration signal quality by decomposing signals into different wavelet coefficients. Coefficients with various time scales and frequencies represent the signals. Then, the signal has to be analyzed frequently, and the wavelet coefficients thresholding values are utilized to remove the noise from the signal. Suppose the signal magnitude value is below the threshold value; set zero as the coefficient value. In addition, the wavelets identify the features in the vibration data by preserving the signal features. After removing the signal coefficients, inverse wavelet transform is applied to get the original input. The gathered vibrated data effectively identifies the faults, severity, and types.

Considered $x(t)$ is denoted as the rolling bearing vibration signal, and it is affected by noise $n(t)$. Then, the wavelet representation of the original vibration signal is defined using equation (1)

$$y(a,b) = \int x(t)\psi^{(a,b)} * t \; dt \tag{1}$$

In equation (1), the vibrate signal wavelet function is defined as $\psi^{(a,b)}$, translation factor is b, and scale factor is a. The $n(t)$ is eliminated from the $x(t)$ were using the thresholding wavelet coefficient process. The threshold value and magnitude of the signal have to be computed for the signal, and the low-value coefficients are set as zero. Using Stein's unbiased risk estimate (SURE) process, the denoising procedure determines the coefficient threshold value. The SURE intention is to reduce the denoised signal risk. The threshold value is estimated using equation (2)

$$\left.\begin{aligned} t_{SURE} &= argmin_t E\left[\left|x - x_{den}(t)\right|^2\right] \\ E\left[\left|x - x_{den}(t)\right|^2\right] &= E\left[x - y(t)^2\right] + E\left[\left|y(t) - x_{den}(t)\right|^2\right] \\ t_{SURE} &= \sqrt{\frac{\log(2) * sigma^2}{alpha}} \end{aligned}\right\} \tag{2}$$

In equation (2), the wavelet transform of vibration signal x is defined as $y(t)$, with a threshold value. The threshold value is estimated by the squared error value between the original and wavelet transform at the t threshold value. The SURE threshold is estimated to reduce the expected square error from the computed value. Then, the expected squared error value from t to zero then the threshold value is defined using equation (2). Here, the $sigma^2$ is denoted as noise variance, and the constant value is denoted as *alpha*, determined depending on the wavelet function. In equation (2), a denoised signal with a threshold value is denoted as $x_{den}(t)$, original vibrate signal is represented as x, and the expectation operator is defined as $E[.]$ The SURE thresholding is robust and minimizes the risk factors successfully. According to the threshold values, the low-magnitude coefficients are removed from the signal. The original signals are then reconstructed by applying the inverse wavelet transform, as defined in equation (3).

$$x_{den}(t) = \Sigma(a,b)\, y(a,b)\, \psi^{(a,b)} * t \tag{3}$$

The equations (1 and 2) are utilized for removing the noise from the vibrate signals, and the original signal is reconstructed based on the inverse wavelet transform defined in equation (3). Then, the overall noise removal process of rolling bearing vibrate data is illustrated in Fig. 2.

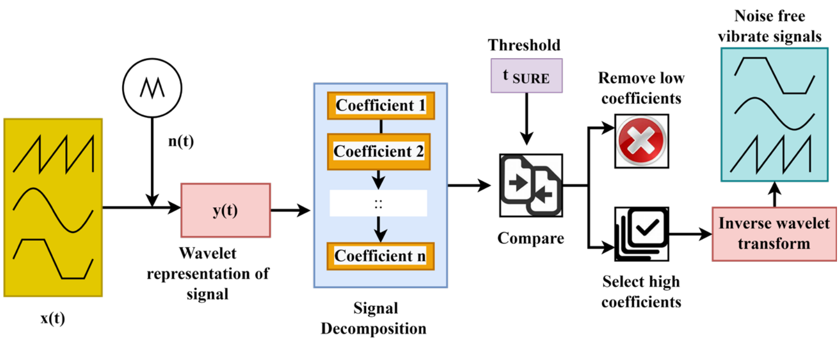


Fig. 2. Process of wavelet denoising-based vibrate signal noise removal

Fig. 2 illustrates that the wavelet denoising process successfully removes the irrelevant and outlier information. A data segmentation process processes the noise-removed bearing signals to improve the fault diagnosis accuracy. Segmentation is splitting the long sequence information into smaller segments with manageable sections. As said, the bearing sensor information is continuously collected to identify the faults in an earlier stage. Therefore, the gathered bearing information's long sequences because it captures the operating conditions, durations, and multiple fault occurrences. The bearing data segmentation process isolates the specific faults during the bearing operations. The isolated features maximize the accurate fault diagnosis in large volumes of data.

The segmentation process reduces difficulties such as time consumption and is expensive in analyzing the large volume of bearing data. After decomposing the information, the application of feature extraction and classification process minimizes the entire computational burden. Initially, segment overlapping and length need to be identified in the segmentation process. The segment length has been decided depending on the bearing system characteristics, requirements, and expected fault durations. After that, overlapping segments are identified due to the segments starting point and shifting process. The dataset has the labeling concept that labels the segmented data as fault and fault-free, which helps to identify the patterns. Predicted patterns are more useful for diagnosing new patterns with maximum accuracy and reducing computation difficulties. After segmenting the data, it has been fed into the feature extraction process to derive the features of the bearing system.

3.3 Feature Extraction

The segmented vibrate signals are fed into the Ant-Optimized Long Short-Term Neural Model (AO-LSTM) to extract the bearing features. Feature extraction is the way of deriving the features from the vibrate signals. The feature extraction process involves different aspects, such as time-frequency, domain, and frequency analysis. This work uses the AO-LSTM approach to derive the bearing features because it can process the time-series data with minimum computation challenges. The neural model can learn the long-term dependencies that highly influence feature extraction. The automatic learning process derives the features from bearing systems for the fault diagnosis classification process. The AO-LSTM can derive a large number of features and is easy to train, that are the main reason for selecting these methods to perform feature extraction. The LSTM is one of the Recurrent Neural Networks with memory to store entire processing time-series data. The LSTM works with the Ant Colony Optimization (ACO) algorithm to fine-tune the network parameters that provide the optimal solutions while extracting features.

The network has four types of gates and LSTM cells that participate in computing the output value. The LSTM cell is the network's computation unit that helps process a single input and ensures the single output value. The cell receives the current time step feature vector and produces the output, utilized in the next cell in LSTM networks. The forget gate determines how much previously processed information has to be forgotten. The input gate determines how much present inputs are included in the cell state, and the output gate determines how much cell state has to be output. The cell state has been updated frequently after completing each gate performance. The network uses the segmented vibrate data as input, which are processed with the help of four gates and produces the features as the output value. Then, the LSTM output computation is defined in equation (4).

$$\left.\begin{aligned} f(t) &= sigmoid\left(W_f * x(t) + U_f * h_{t-1}\right) \\ i(t) &= sigmoid\left(W_i * x(t) + U_i * h_{t-1}\right) \\ c_t &= f(t) * c_{t-1} + i(t) * \tanh\left(W_c * x(t) + U_c * h_{t-1}\right) \\ o(t) &= sigmoid\left(W_o * x(t) + U_o * h_{t-1}\right) \\ h(t) &= o(t) * \tanh(c_t) \end{aligned}\right\} \quad (4)$$

In equation (4), the output is obtained in the output gate at t time $o(t)$ by considering the forget gate process $f(t)$, input gate $i(t)$, hidden state $h(t)$ and cell state c_t at time step t. The computation uses the input vector $x(t)$ at time t. For every computation, the network verifies the parameters, such as weight metrics (W and U). The output is computed with the help of the tanh and sigmoid activation functions, which maximizes the feature extraction efficiency. The rolling bearing fault diagnosis process requires numerous features such as vibration amplitude, frequency, frequency spectrum, time domain, and frequency domain features. These features are widely utilized to recognize the faults in bearings. The faults are identified from the increased vibration amplitude, frequency shift, frequency changes, and time domain representations.

The LSTM uses the four gates and sigmoid activation function to produce the output values such as 0 and 1. The outputs belong to the event happening in bearing systems. During the output estimation, LSTM uses back-propagation through time (BPTT) to train the vibrate data by considering the temporal dependencies of the data. The LSTM receives the segmented vibrate signal as input in which signals are decomposed into the sequence of time steps). Then, LSTM is trained depending on the time sequences used to identify the vibration signal at the time step. As said, the LSTM has learning patterns from the long-term dependency information, which is used to identify the vibration with the maximum prediction rate. Then, the vibration amplitude computation is defined in equation (5)

$$z(t) = f\left(x(t), h_{t-1}\right) \quad (5)$$

In equation (5), vibration amplitude is denoted as $z(t)$ that is computed from the input signal $x(t)$, and previous time step hidden state information. Let's assume the vibrating signal is split into the sequence of 100-time steps, and the LSTM is applied to the time-series data to predict the vibrate signal amplitude value. The network uses the BPTT training process that maximizes the amplitude computation efficiency. Therefore, the Peak-to-Peak amplitude value is extracted as a feature in vibration amplitude. After that, frequency information is extracted from the vibrating signal using LSTM. The frequency measures how often the signal is repeated, measured in terms of hertz (HZ) or cps (Cycle Per Second). LSTM uses the Fourier Transform to decompose the signals into constituent frequencies. The transform examines the frequency components with the help of the network learning process. Considering the bearing machines suddenly raise their sound, they face failure issues. Therefore, the LSTM approach observes the vibration signal frequency to prevent machine failure. The frequency spectrum features are identified by considering the vibration signal and its amplitude. Then, statistical computations like standard deviation, mean, kurtosis, and skewness values are extracted. In addition, spectral spread, spectral centroid, and spectral kurtosis information are extracted. The derived features are utilized to classify the faults in bearing machines. During the feature extraction process, network performance is improved by optimizing their parameters. The optimization is done with the help of the Ant Colony Optimization (ACO) algorithm. The ACO algorithm generates the ant populations and initializes the pheromone level as 1 for all actions. The ants or feature having the actions in the initial state and the loss value is estimated for every feature extraction. If the deviations are high, parameters like pheromone level are updated to minimize the loss value. Then, the updating process on LSTM is defined using equation (6).

$$p\left(a_t|a_{t-1}\right)=\frac{\tau_t^{a_t}}{\sum_{a\in A}\tau_t^{a_t}} \tag{6}$$

In equation (6), the probability of action a_t at t time step is $p(a_t|a_{t-1})$ and the previous action is defined as a_{t-1}, pheromone level of action [a] at t is denoted as τ_t^a, and set of possible actions is denoted as A. According to equation (6), the parameters are computed frequently, and the network is fine-tuned to reduce the deviations between the outputs. Then, the overall process of the AO-LSTM process is defined in Fig. 3.

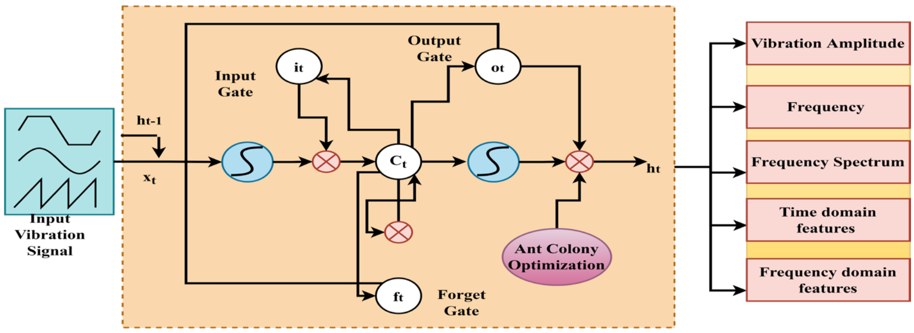


Fig. 3. Process of AC-LSTM-based vibration signal feature extraction

Fig. 3 illustrates the process of the AC-LSTM-based vibration feature extraction process. The neural model derives the vibrate features such as vibration amplitude, frequency spectrum, time, and frequency domain features. These features are widely utilized to classify bearing faults at various locations. Among the various features, vibration amplitude, shift frequency, time, and frequency domain features are widely used to identify faults at various locations and operations. The extracted features are fed into the Autoencoder Neural Model (ANM) to diagnose the faults.

3.4 Fault Diagnosis using Autoencoder Neural Model (ANM)

The final step of this work is diagnosing the faults from rolling bearing machines, done with the help of the Autoencoder Neural Model (ANM). The network is an unsupervised learning approach that predicts the output by mapping the inputs with feature space. The working process of ANM is illustrated in Fig. 4.

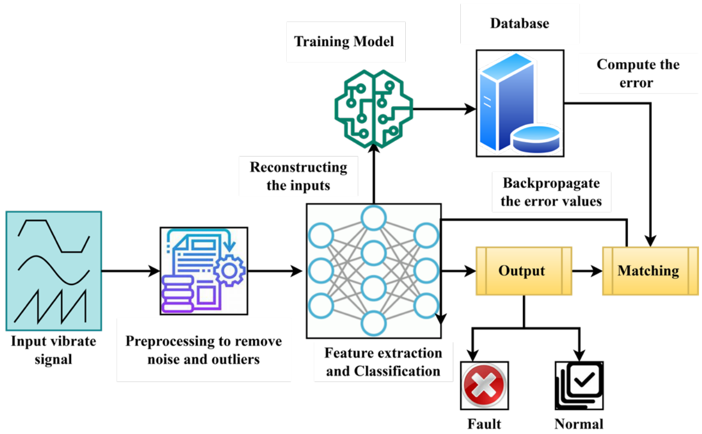


Fig. 4. Autoencoder neural model-based fault diagnosis identification process

The network has an encoder and decoder; the encoder process maps the inputs f into the low-dimensional feature space using the function $h = f(x)$. Then the decoder g gets the low-dimensional feature again, mapping into the input with the feature space, and the mapping process is defined as $\hat{x} = g(h)$. The feature mapping and learning process computes the outputs, which helps to recognize the faults. During the classification, this study intends to reduce the loss function, which is computed using equation (7)

$$L = \frac{1}{n}\sum_{i=1}^{n}\left\| x_i - Net\left(x_i\right)\right\| \tag{7}$$

In equation (7), L is defined as the loss function that is computed from the Euclidean distance between the x and $\hat{x}$. The distance measure identifies the error rate; according to the error value, the network parameters are fine-tuned to maximize the fault detection accuracy. The encoder output h is the low-dimensional representation of the input. The network has multiple hidden layers that process the input to identify the output values with maximum accuracy. The hidden layers are utilized in the inputs to reconstruct the input signal, which helps to predict the fault. The autoencoder process is applied to the input features, and a learning process is initiated to predict the faults at various locations and operations. The trained templates are more useful to identify whether the vibrate signals are normal or faults. The autoencoder process can understand and learn the nonlinear complex relationship between the fault conditions and vibration signals. The calculated reconstruction error values identify the exact faults with minimum error and maximize accuracy. Then, the system's overall efficiency is evaluated using experimental results and discussions.

4 Results and Analysis

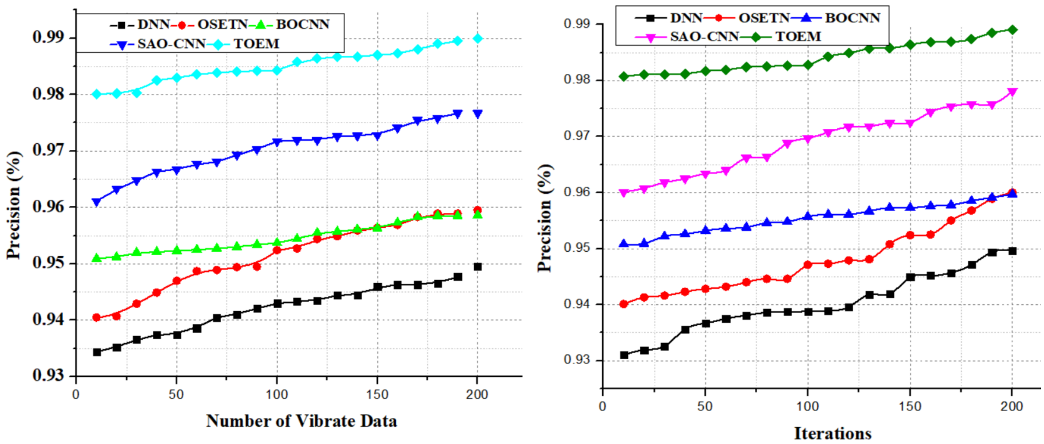
This section discusses the efficiency of the Triplet Optimized Embedding Model (TOEM)-based fault diagnosis process. The triplet process uses the wavelet denoising technique to remove the noise from the vibrating signal. During the analysis, signals are decomposed according to the wavelet transform that identifies every small deviation in the vibrating signal. The SURE threshold values and the wavelet denoising process are worked to identify the low magnitude coefficient values. The computed low coefficient values are removed from the vibrate signals, which maximizes the fault diagnosis accuracy. In addition, long-short-term memory networks frequently analyze the time-series data that identifies the vibrating signal features with maximum accuracy. The data segmentation process, which divides the large volume of data that simplifies the computation difficulties, overcomes the high-dimensionality issues. The autoencoder networks reconstruct the input signals according to the encoder and decoder process that effectively categorizes the normal and fault signals. The introduced TOEM system uses the Case Western Reserve University bearing center information to explore the system's efficiency. The TOEM approach efficiency is described with metrics such as True Positive Rate (TPR), False Positive Rate

(FPR), Precision, Recall, Time to detection, and F1-score. Then, the TOEM efficiency is compared with existing methods, such as Deep Neural Networks (DNN) [18], Optimal Ensemble Deep Transfer Networks (OSETN) [19], Bayesian Optimization Convolution Neural Networks (BOCNN) [20], and Simulated Annealing Optimized Convolution Neural Networks (SAO-CNN) [21]. Then, the results of TPR and FPR are illustrated in Table 2. The TPR is measured as how the TEOM correctly identifies the faults in the system. The FPR is measured as how the TEOM incorrectly identifies the faults in the system.

Table 2. TPR vs. FPR analysis of TOEM

Methods	True Positive Rate (TPR)	False Positive Rate (FPR)
DNN [18]	0.9	0.1
OSETN [19]	0.82	0.18
BOCNN [20]	0.678	0.322
SAO-CNN [21]	0.468	0.532
TOEM	**0.367**	**0.633**

Table 2 illustrates that the DNN has the highest TPR (0.932) and high FPR values, which means the DNN method is sensitive to faults. Therefore, the DNN having the high incorrectly identifies the faults. Compared to other methods, the TOEM approach attains the minimum TPR values and the same as FPR values, indicating that the TOEM method has minimum incorrectly identified the faults in the bearing systems. The vibrate signal has been preprocessed with the help of the wavelet denoising process that removes the low-magnitude information from the original signal. Properly analyzing the input signals is more useful in reducing incorrect fault identification.

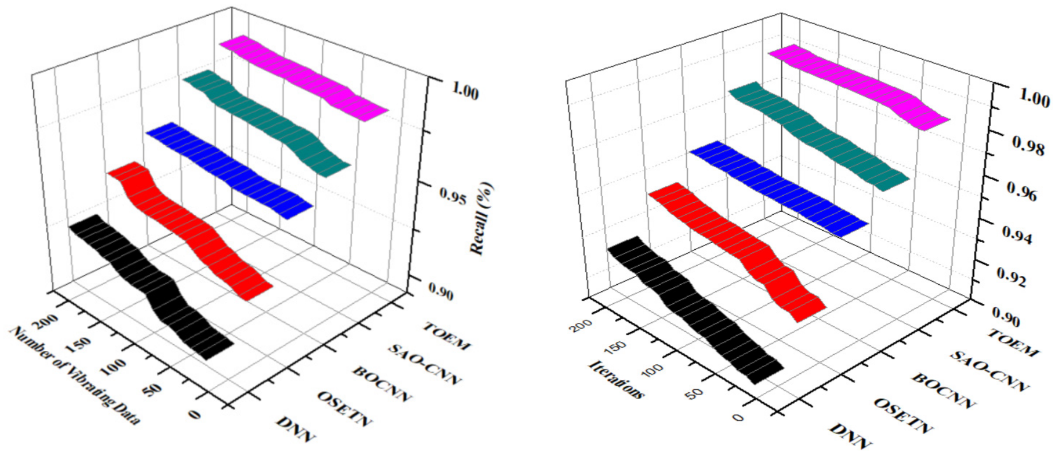


(a) Vibrate data (b) Iterations

Fig. 5. Precision analysis for TOEM

Fig. 5 illustrates the efficiency of the TOEM approach-based fault diagnosis process. The efficiency is analyzed for different numbers of data and iterations. The TOEM approach uses triplet methods such as wavelet denoising, and ant colony-optimized long short-term networks, and an autoencoder neural model. The wavelet denoising process uses the SURE threshold value that identifies the signal magnitude, which helps to remove the outlier data. In addition, the autoencoder model reconstructs the inputs and generates the template for every input. Then, the deconstruction process is applied to get the output process. The encoder model performs the mapping process that identifies fault features from the vibrating amplitude, time domain, and frequency information. Then, TOEM's precision analysis is measured against that of well-established techniques like Deep Neural Networks (DNN) [18], Optimal Ensemble Deep Transfer Networks (OSETN) [19], Bayesian Optimization Convolution Neural Networks (BOCNN) [20], and Simulated Annealing Optimized Convolution Neural Networks (SAO-CNN) [21].

The effective utilization of the network process increases the fault recognition rate for various numbers of data (98.95%) and iterations (98.47%). The obtained precision value is higher than the other methods.



(a) Vibrate data (b) Iterations

Fig. 6. Recall analysis for TOEM

Fig. 6 shows the recall analysis of TOEM based fault recognition process in rolling bearing Machinery. The ant colony-optimized long short-term neural model extracts features from vibrating signals. The extracted vibrating amplitude, feature domain, and time domain information resemble the fault information in vibrating signals. During the analysis, the Fourier transform is frequently applied to the vibrating signals that segment the signals, improving the overall fault recognition rate. The extracted features are fed as the encoder neural model input, reconstructing the inputs and mapping with the feature search space. The Recall analysis of TOEM is then evaluated in comparison to other popular approaches, including Deep Neural Networks (DNN) [18], Optimal Ensemble Deep Transfer Networks (OSETN) [19], Bayesian Optimization Convolution Neural Networks (BOCNN) [20], and Simulated Annealing Optimized Convolution Neural Networks (SAO-CNN) [21].

This mapping process identifies the faults features with minimum deviation error rate. The computed output value is compared with the template features using Euclidean distance measures that identify the output error. The computed error values are corrected with the help of the backpropagation learning process, which enhances the accuracy of the overall fault diagnosis. In addition, the overall analysis of the TOEM method efficiency is described in Table 3.

Table 3. F1-Score analysis

Vibrating data	DNN [18]	OSETN [19]	BOCNN [20]	SAO-CNN [21]	TOEM
10	0.9427	0.9417	0.9413	0.9649	**0.9782**
20	0.9435	0.9445	0.9573	0.9617	**0.9867**
30	0.9415	0.9317	0.9339	0.9432	**0.9803**
40	0.938	0.9505	0.9298	0.9501	**0.9707**
50	0.9337	0.9368	0.9534	0.9624	**0.9854**
60	0.9313	0.943	0.9249	0.9484	**0.9712**
70	0.9373	0.9459	0.9547	0.9589	**0.9796**
80	0.9458	0.9408	0.9586	0.9668	**0.9718**
90	0.9412	0.9582	0.9494	0.9421	**0.9862**
100	0.949	0.9563	0.9466	0.9613	**0.985**
110	0.9395	0.9533	0.9398	0.9552	**0.9831**
120	0.9318	0.9502	0.9217	0.9666	**0.9864**

130	0.9347	0.9519	0.9444	0.9514	**0.9759a**
140	0.9398	0.9421	0.9282	0.9425	**0.9896**
150	0.9317	0.9521	0.94	0.9602	**0.9745**
160	0.9485	0.9512	0.9454	0.9622	**0.9807**
170	0.946	0.9433	0.9532	0.9687	**0.9805**
180	0.9361	0.9332	0.9348	0.9661	**0.9742**
190	0.9463	0.9386	0.9414	0.9418	**0.9832**
200	0.9352	0.9423	0.9539	0.9659	**0.9801**

Table 3 analyzes the F1-Score value of the TOEM fault diagnosis process in which the proposed TOEM attains maximum fault diagnosis value compared to other methods. The TOEM attains 97.85% accuracy, which is higher than other classifiers such as Deep Neural Networks (DNN) [18], Optimal Ensemble Deep Transfer Networks (OSETN) [19], Bayesian Optimization Convolution Neural Networks (BOCNN) [20], and Simulated Annealing Optimized Convolution Neural Networks (SAO-CNN) [21]. The method explores every vibrating signal using a wavelet denoising procedure that removes irrelevant information and minimizes computation difficulties. In addition, the sequence time-series information widely stores every processing input that helps identify the new patterns with minimum deviation errors. Moreover, the method consumes minimum computation time for predicting the faults in rolling bearing machines. The obtained time value is shown in Table 4.

Table 4. Time analysis

Vibrating data	DNN [18]	OSETN [19]	BOCNN [20]	SAO-CNN [21]	TOEM
10	0.2012	0.202	0.1829	0.1776	**0.1075**
20	0.2247	0.1775	0.1442	0.125	**0.1035**
30	0.2142	0.1607	0.1422	0.134	**0.125**
40	0.2162	0.1986	0.1479	0.1333	**0.1071**
50	0.1925	0.1657	0.1612	0.1672	**0.1022**
60	0.1942	0.1747	0.1627	0.1274	**0.1051**
70	0.2255	0.1832	0.1558	0.121	**0.1046**
80	0.2072	0.1612	0.168	0.1507	**0.122**
90	0.2209	0.1874	0.1862	0.127	**0.103**
100	0.1915	0.1895	0.1899	0.1384	**0.124**

Table 4 illustrates the time analysis for the fault diagnosis process in rolling bearing machines. The method attains minimum computation time (0.11s) compared to other methods; if the number of data increases, then computation time also increases. The TOEM approach recognizes the faults by embedding the different methods on vibrating signals to improve the overall fault diagnosis efficiency in high-dimensional space. During the analysis, the autoencoder neural model predicts the error rate that identifies the deviations between the predicted and computed output value. The input reconstruction and memory network-based extracted features are more useful for identifying faulty and healthy bearings. The network uses the embedding, which uses the wavelet denoising and SURE threshold values that remove the low coefficients effectively. Then, the encoder network trains the systems to improve the overall recognition rate with minimum time and loss. Therefore, the computed loss function and output computation improves the overall efficiency; however, it depends on the dataset size. Thus, the TOEM method recognizes the bearing faults with maximum recognition accuracy (97.85%) and minimum computation time (0.11s).

5 Conclusion

Thus, the paper analyzes the Triplet Optimized Embedding Model (TOEM) to diagnose the faults from the rolling bearing machines. The vibrating signals are initially collected from Case Western Reserve University bearing center to diagnose the faults. The gathered vibrate information is processed using wavelet denoising that removes the irrelevant and outlier information by comparing the signal threshold value with the coefficients. This process removes the low coefficient values, and a data segmentation process is performed to decompose the large-scale dataset into subsets. Then, a long-term, short-term memory network extracts the features. The derived features are representations of the faults in different locations and operations. The feature extraction process is further improved by optimizing the network with ant colony optimization. The extracted features are fed into the auto-encoder neural model that reconstructs the inputs, and a mapping process is performed to identify the faults with high recognition accuracy (97.85%). In the future, the fault diagnosis process will be enhanced by selecting the optimized features that are performed with the help of the feature selection process.

Acknowledgements

This research was supported by "Research on Intelligent Motor Bearing Failure Monitoring System" (No. 2021KY1405), "Research on Intelligent Monitoring and Early-Warning System for Rutting Machine Notch" (No. 2022KY1422), "Research on Thousand Young and Middle-aged Backbone Teachers Cultivation Programme of Guangxi Higher Education Institutions in 2020", and "Research on Intelligent Monitoring and Warning System of Rutting Machine Notch" (No. 2022KY1422), "Research on Thousand Young and Middle-aged Backbone Teachers Cultivation Programme of Guangxi Higher Education Institutions in 2020" (approved by "Teachers in Guangxi Teaching NO.[2020]58").

References

[1] S. Liu, H. Jiang, Z. Wu, X. Li, Data synthesis using deep feature enhanced generative adversarial networks for rolling bearing imbalanced fault diagnosis, Mechanical Systems and Signal Processing 163(2022) 108139.
[2] M.-A. Khan, B. Asad, K. Kudelina, T. Vaimann, A. Kallaste, The Bearing Faults Detection Methods for Electrical Machines—The State of the Art, Energies 16(1)(2023) 296.
[3] A. Choudhary, T. Mian, S. Fatima, Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images, Measurement 176(2021) 109196.
[4] R.-N. Toma, A.-E. Prosvirin, J.-M. Kim, Bearing fault diagnosis of induction motors using a genetic algorithm and machine learning classifiers, Sensors 20(7)(2020) 1884.
[5] T. Mian, A. Choudhary, S. Fatima, A sensor fusion based approach for bearing fault diagnosis of rotating machine, Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability 236(5)(2022) 661-675
[6] O. Janssens, M. Loccufier, S.-V. Hoecke, Thermal imaging and vibration-based multisensor fault detection for rotating machinery, IEEE transactions on industrial informatics 15(1)(2019) 434-444.
[7] K. Kaplan, Y. Kaya, M. Kuncan, M.-R. Minaz, H.-M. Ertunç, An improved feature extraction method using texture analysis with LBP for bearing fault diagnosis, Applied Soft Computing 87(2020) 106019.
[8] L. Eren, Bearing fault detection by one-dimensional convolutional neural networks, Mathematical Problems in Engineering 2017(2017) 1-9.
[9] H.-Y. Pan, H.-F. Xu, J.-D. Zheng, J. Su, J.-Y. Tong, Multi-class fuzzy support matrix machine for classification in roller bearing fault diagnosis, Advanced Engineering Informatics 51(2022) 101445.
[10] J.-H. Zhong, P.-K. Wong, Z.-X. Yang, Simultaneous-fault diagnosis of gearboxes using probabilistic committee machine, Sensors 16(2)(2016) 185.
[11] K.-I. Edomwandekhoe, Modeling and fault diagnosis of broken rotor bar faults in induction motors, [PhD dissertation] Newfoundland and Labrador: Memorial University of Newfoundland, 2018.
[12] M. Jung, O. Niculita, Z. Skaf, Comparison of different classification algorithms for fault detection and fault isolation in complex systems, Procedia Manufacturing 19(2018) 111-118.
[13] D.-Y. Dou, S.-H. Zhou, Comparison of four direct classification methods for intelligent fault diagnosis of rotating machinery, Applied Soft Computing 46(2016) 459-468.
[14] P. Chen, X.-Q. Zhao, Q.-X. Zhu, A novel classification method based on ICGOA-KELM for fault diagnosis of rolling bearing, Applied Intelligence 50(9)(2020) 2833-2847.
[15] J. Pacheco-Chérrez, J.-A. Fortoul-Díaz, F.Cortés-Santacruz, L.-M. Aloso-Valerdi, D.-I. Ibarra-Zarate, Bearing fault detection with vibration and acoustic signals: Comparison among different machine leaning classification methods, Engineering Failure Analysis 139(2022) 106515.
[16] L.-H. Yuan, D.-S. Lian, X. Kang, Y.-Q. Chen, K.-J. Zhai, Rolling bearing fault diagnosis based on convolutional neural network and support vector machine, IEEE Access 8(2020) 137395-137406.

[17] J.-M. Li, X.-F. Yao, X.-D. Wang, Q.-W. Yu, Y.-G. Zhang, Multiscale local features learning based on BP neural network for rolling bearing intelligent fault diagnosis, Measurement 153(2020) 107419.
[18] C.-J. Wang, Z.-L. Xu, An intelligent fault diagnosis model based on deep neural network for few-shot fault diagnosis, Neurocomputing 456(2021) 550-562.
[19] X.-Q. Li, H.-K. Jiang, R.-X. Wang, M.-G. Niu, Rolling bearing fault diagnosis using optimal ensemble deep transfer network, Knowledge-Based Systems 213(2021) 106695.
[20] D. Kolar, D. Lisjak, M. Pająk, M. Gudlin, Intelligent fault diagnosis of rotary machinery by convolutional neural network with automatic hyper-parameters tuning using Bayesian optimization, Sensors 21(7)(2021) 2411.
[21] F. He, Q. Ye, A bearing fault diagnosis method based on wavelet packet transform and convolutional neural network optimized by simulated annealing algorithm, Sensors 22(4)(2022) 1410.
[22] S.-Z. Gao, L.-T. Xu, Y.-M. Zhang, Z.-M. Pei, Rolling bearing fault diagnosis based on SSA optimized self-adaptive DBN, ISA transactions 128(2022) 485-502.
[23] D.-C. Yao, H.-C. Liu, J.-W. Yang, X. Li, A lightweight neural network with strong robustness for bearing fault diagnosis, Measurement 159(2020) 107756.
[24] R.-X. Wang, H.-K. Jiang, X.-Q. Li, S.-W. Liu, A reinforcement neural architecture search method for rolling bearing fault diagnosis, Measurement 154(2020) 107417.
[25] Y. Guan, M. Zong, D.-Y. Sun, J.-B. Liu, F.-J. Fan, 2MNet: Multi-sensor and multi-scale model toward accurate fault diagnosis of rolling bearing, Reliability Engineering & System Safety 216(2021) 108017.
[26] R.-N. Liu, B.-Y. Yang, A.-G. Hauptmann, Simultaneous bearing fault recognition and remaining useful life prediction using joint-loss convolutional neural network, IEEE Transactions on industrial informatics 16(1)(2020) 87-96.
[27] X.-D. Song, Y.-Y. Cong, Y.-F. Song, Y.-L. Chen, P. Liang, A bearing fault diagnosis model based on CNN with wide convolution kernels, Journal of Ambient Intelligence and Humanized Computing 13(8)(2022) 4041-4056.
[28] L. Leng, J. Zhang, Palmhash code vs. palmphasor code, Neurocomputing 108(2013) 1-12.
[29] L. Leng, M. Li, C. Kim, X. Bi, X, Dual-source discrimination power analysis for multi-instance contactless palmprint recognition, Multimedia tools and applications 76(2017) 333-354.
[30] L. Leng, M. Li, A.B.J. Teoh, Conjugate 2DPalmHash code for secure palm-print-vein verification, in: Proc. 2013 6th International congress on image and signal processing (CISP), 2013.
[31] Case Western Reserve University, Bearing Data Center. <https://engineering.case.edu/bearingdatacenter>, (accessed 24.08.2023).