ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus
Mécanique

Miloud Sedira and Ahmed Felkaoui

A hybrid HMM-RBFNN system using relevant and non-redundant features for bearing fault classification

Volume 354 (2026), p. 365-395

Online since: 15 April 2026

https://doi.org/10.5802/crmeca.350

The Comptes Rendus. Mécanique *are a member of the Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1873-7234

ACADÉMIE DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus. Mécanique
2026, Vol. 354, p. 365-395
https://doi.org/10.5802/crmeca.350

Research article

A hybrid HMM-RBFNN system using relevant and non-redundant features for bearing fault classification

Miloud Sedira [a,*] **and Ahmed Felkaoui** [a]

[a] LMPA, Ferhat Abbas University, Setif 19000, Algeria

E-mails: miloudsedira@yahoo.fr (M. Sedira), a_felkaoui@yahoo.fr (A. Felkaoui)

Abstract. This paper presents an innovative hybrid approach that combines Hidden Markov Models (HMM) with Radial Basis Function Neural Networks (RBFNN) for the automatic classification of mechanical faults in rolling element bearings using vibration signal analysis. The signals are sourced from the well-established database (https://engineering.case.edu/bearingdatacenter/welcome), widely used in fault diagnosis research. Raw signals are preprocessed to extract relevant features across time, frequency, and time–frequency domains, including wavelet packet decomposition. To enhance classification robustness and reduce computational complexity, dimensionality reduction is performed using Principal Component Analysis (PCA), complemented by Fisher score-based feature selection. HMMs are trained to capture the temporal dynamics of the signals, while RBFNNs leverage the reduced feature space for fine-grained classification. A comprehensive performance comparison is conducted between standalone HMM and RBFNN models, as well as their integration within the hybrid HMM-RBFNN system. Experimental results demonstrate that the proposed hybrid method significantly improves classification accuracy, highlighting its potential for industrial predictive maintenance applications.

Keywords. Bearing, Diagnosis, Fisher score, Hidden Markov model (HMM), Principal component analysis (PCA), Radial basis function neural network (RBF), Wavelet packet transform (WPT).

Note. Submitted by invitation following the DTE-AICOMAS 2025 conference, held February 17–21, 2025.

Manuscript received 25 November 2025, revised 19 January 2026 and 18 March 2026, accepted 12 February 2026, online since 15 April 2026.

1. Introduction and related work

Rotating machines play a key role in industrial production. Once a failure occurs, it can lead to significant losses in time and then in costs [1,2]. Often, these failures are caused by the deterioration of bearings, which are considered one of the most common and vulnerable components of rotating machines [3][1]. Indeed, many rotating machines use bearings, including aircraft engines, high-speed trains, helicopters, wind turbines, machine tools, industrial robots, etc. It is significant that bearings are responsible for 30% or more of the overall failures of rotating machines [4,5]. Indeed, condition monitoring, diagnosis, and identification of bearing faults constitute an important guarantee to improve the reliability of rotating machines and industrial equipment [6]. For over two decades, great importance has been given to research on

* Corresponding author

[1] https://engineering.case.edu/bearingdatacenter/welcome.

methodologies for automating the diagnosis of mechanical faults in rotating machines, based on the analysis of vibration signals, acquired using well-suited sensors installed near the mechanical elements to be monitored [7,8]. Indeed, fault detection and diagnosis can be considered a pattern recognition problem related to the state of rotating equipment; they can be divided into three phases: data acquisition, feature extraction, and fault classification, with the latter two being priorities [9,10]. Given that the information contained in vibratory signals is complex and variable, it is important for diagnosis to accurately extract the intrinsic characteristics of all fault types [11]. The actually collected signal often contains noise, and only by suppressing this noise can the useful information in the original signal be represented and exploited effectively [1,12]. Signal processing is one of the most commonly used methods for the second stage, typically using time-domain, frequency domain, or time–frequency analysis [5]. Regarding classification, different algorithms can then be used to classify the faults according to the predefined classes [10]. The techniques for diagnosing mechanical faults in rotating machines currently available present various limitations. A robust and effective method must be sought, and an automated system must be developed for diagnostic activities [13]. Indeed, recently, the intelligent fault diagnosis approach is widely adopted in rotating machine condition monitoring systems, especially with the advent of Industry 4.0 technologies. Throughout the literature, we notice that the second stage plays a more important role; it aims to extract representative features from the original signals [14]. Similarly, many studies have focused on how to extract relevant and effective features from measured signals in the time domain, based on various signal processing techniques [12]. In general, time-domain analysis calculates statistical parameters such as root mean square (RMS), kurtosis, structural resonances, etc. Frequency domain analysis is often advantageous, as it allows for easy isolation and identification of key frequency components. A commonly used tool is the Fast Fourier Transform (FFT), as well as FFT-based methods, spectral analysis methods, etc. Time–frequency analysis extends the capabilities of frequency analysis to non-stationary vibration signals and includes methods such as the Short-Time Fourier Transform, Wavelet Transform, Empirical Mode Decomposition (EMD), Hilbert-Huang Transform (HHT), methods associated with EMD, etc. With the development of nonlinear dynamic theory, many entropy-based estimation methods offer useful alternative approaches for extracting features related to faults hidden in vibration signals and applying them to bearing fault detection [6]. Features extracted from vibration signals may include redundant or insensitive information; therefore, some dimensionality reduction strategies are applied, such as Principal Component Analysis (PCA) [15], Linear Discriminant Analysis (LDA), the distance evaluation technique, and Kullback–Leibler (K–L) divergence, or class separation tools such as the Fisher Score (FS). As for classification methods, they implement a mapping of the fault feature vector using as inputs the features or descriptors established in the second stage. With the dizzying development of computing means, this aspect has seen rapid advancement; indeed, researchers and engineers in the field have striven to implement a plethora of algorithms and methods ranging from statistical/probabilistic methods such as Gaussian Mixture Models (GMM), Bayesian inferences, Hidden Markov Models (HMM) to artificial intelligence methods like Multilayer Perceptron (MLP), Radial Basis Function Neural Networks (RBFNN), Support Vector Machines (SVM), Deep Learning (DL), etc.

To support what has just been said, we can cite some relevant landmark work from the last two decades; namely that [5] implemented a procedure for diagnosing bearing faults using three classifiers, namely Multilayer Perceptron (MLP), Radial Basis Function Neural Networks (RBFNN), and Probabilistic Neural Networks (PNN). Feature selection and optimization were performed by Genetic Algorithms (GA) from statistical data extracted from vibration signals in the time domain; appreciable results were recorded [16] used statistical features extracted from the time domain as input data for a feedforward neural network for bearing fault classification [2] demonstrated

the performance of SVM compared to vector quantization and self-organizing maps, after a comparison performed for bearing fault diagnosis. Fault features were extracted from vibration signals by wavelets, and for the selection of the most relevant features, a statistical methodology based on the minimum Shannon entropy criterion was employed [17] conducted a comparison between RBFNN and traditional ANN. Indeed, an application of 4 practical examples was implemented. Following this comparison, they concluded that RBFNN is specially recommended for function approximation problems, while for classification problems, traditional neural networks can achieve better results with more efficiency for RBF concerning learning when the number of data is optimal [18] presented a method for diagnosing bearing faults from multiple sources. Thus, Multiple Frequency Energy Spectrum (MFES) was used for feature extraction, to improve recognition accuracy and reduce uncertainty. Fault classification was performed by an RBFNN [19] showed after a survey on RBFNN that they have an appreciable generalization capacity and are recommended for the approximation and classification of nonlinear functions; which constitutes a good alternative to MLP [20] used the Continuous Wavelet Transform (CWT) (Meyer wavelet) for extracting bearing fault features and SVM for the classification task. Results achieved 100% accuracy [21] presented a bearing fault diagnosis system based on MLP. They introduced a vibration feature selection based on automatic Bayesian relevance determination, knowing that they used multiple data sources. MLP were dedicated to the classification task [7] defined a new approach to characterize bearing degradations and simplify the prognostic task into a classification task. Indeed, they used the Weibull distribution for feature selection and artificial neural networks for fault classification; an application on bearings was performed to validate this proposal [11] proposed a feature fusion model based on a Probabilistic Neural Network (PNN) and entropy. Thus, three types of entropy were extracted from vibration signals, in the time domain, frequency domain, and time–frequency domain. These entropies served as inputs to the classification system based on a PNN network [15] proposed a bearing diagnosis system based on SVM and PCA. Feature extraction from vibration signals was performed by Wavelet Packet Transform (WPT), feature selection by PCA, and classification by SVM and ANN of two types (BP and RBFNN). They concluded that PCA effectively eliminates redundant features, which ensures the performance and reliability of SVM for fault classification relative to the two types of ANN [22] presented a method for diagnosing bearing faults based on Convolutional Neural Networks (CNN); feature extraction from the vibration signal in the time–frequency domain was executed by Wavelet Packet Transform (WPT), then transformed into a 2-dimensional space (image), to subsequently serve as inputs to the CNN for fault classification [13] extracted fault features from bearing vibration signals by Wavelet Packet Energy Entropy (WPEE) to form a feature vector, in order to subsequently serve as inputs to the classification and severity level identification system, namely the multi-class Relevance Vector Machine (mRVM) [23] performed a comparison between Artificial Neural Networks (MLP) and Hidden Markov Models (HMM) for bearing fault classification, using time-domain, frequency-domain, and time–frequency (wavelets) features; indeed, they demonstrated that the performance of MLP diminishes with the increase in the number of features, while HMM remain very robust, which revealed an important performance [13] implemented a technique for diagnosing bearing faults under variable speed conditions, via an improved demodulation spectrum technique using wavelet packets for feature extraction and the Self-Organizing Map (SOM) method for dimensionality reduction [OO] opted for a new approach for the diagnosis and prognosis of bearing faults, based on a multi-feature Hidden Markov Model (HMM). First, the time domain, frequency domain, and wavelet packet decomposition are used to extract state features from bearing vibration signals. The PCA method is then used to reduce their dimensionality. The results reveal that the established approach allows for effective fault diagnosis and estimation of the bearing's remaining useful life [24] proposed a methodology for diagnosing bearing and gear faults, capable of identifying nine different

health state categories; healthy and faulty, under variable and noisy load. A deep neural network is used for fault classification into various categories. Robust features such as semi-variance, spectral kurtosis, and Shannon entropy were used [25] presented a DC motor diagnosis system based on vibration signals; three health states were selected (healthy, incipient fault, and severe fault). Feature extraction was performed by Continuous Wavelet Transform (CWT), dimensionality reduction was made by PCA, while fault classification was performed by K-Nearest Neighbors (KNN) [26] presented a method for diagnosing bearing faults, based on a Convolutional Neural Network (CNN). Feature extraction from the vibration signal in the time–frequency domain was performed by wavelets, then a transformation to a 2-dimensional space (image) was also performed to compose the dataset of inputs to the fault classifier based on Convolutional Neural Networks (CNN).

Recently, many researchers and engineers have turned to the hybrid classification approach using two different classification methods to improve diagnostic performance. In this context, we can cite the work of [12] who developed a hybrid system using LR-type fuzzy logic and ANN to diagnose bearing faults. Wavelet decomposition was employed for feature extraction from vibration signals emitted by the bearings in question [27] presented a method for classifying ball bearing faults based on an FMM-RF hybridization technique, composed of the FMM (Fuzzy Min-Max) neural network and the RF (Random Forest) model. They tested the model on real data, consisting of power spectrum and entropy features. A performance rate of 99.81% was achieved [28] developed a hybrid GHMM-CM method using reduced decomposition features for the recognition and classification of defect types and severity levels. The vibration signal containing defect information was decomposed into several modal components using the VMD method, in which the generalized balancing parameter provides a concise representation of random and epistemic uncertainties. The PCA technique was applied to reduce the dimensionality of the features. Experimental results show that the proposed hybrid GHMM-CM method is more accurate and reliable. Validation of this algorithm was performed on the database (see Footnote 1), Ref. [29] proposed a novel hybrid deep learning method (NHDLM), based on extended deep convolutional neural networks with wide first-layer kernels (EWDCNN) and Long Short-Term Memory (LSTM), for complex environments, which improved the performance of feature classification and offered the best performance and identification accuracy for diagnosing rotating machinery failures [5] implemented a new method for diagnosing bearing faults based on a hybrid model of Convolutional Neural Networks (CNN) and Multilayer Perceptron (MLP). This model simultaneously processes input data of different types and consists of two blocks: MLP for processing numerical inputs and CNN for processing HHT images. They demonstrated that the proposed hybrid model is superior to CNN and MLP models separately.

In conclusion, we find that the hybrid classification approach is more performant than a classifier used separately, that RBFNN, due to the simplicity of their topological structure, their elegant mathematical formulation anchored in classical functional analysis, have proven their rapid convergence, their nature of convex optimizations involved, thus they possess an appreciable generalization capacity, and are highly recommended for the approximation and classification of nonlinear functions [11,21] and that HMM are very robust for modeling time series [3]. This has motivated us to make our contribution, which consists of implementing a hybrid HMM-RBFNN system for the classification of bearing faults.

The originality of our contribution lies in the hybrid system that combines the two previous approaches by exploiting the log-likelihoods from the HMM as additional descriptors to feed the RBFNN. This strategy aims to improve the robustness of the classifier by integrating dynamic information (from HMM) with the statistical features from PCA. To ensure a fair evaluation, all models were trained and tested on the dataset (see Footnote 1), which has become a reference in the field for testing bearing diagnosis algorithms.

The remainder of this document is described as follows: In Section 2, we briefly recall the theoretical foundations of the tools used, namely, statistical tools, frequency and time–frequency analysis tools, RBFNN, and HMM. In Section 3, we present the experimental data used. Section 4 is dedicated to the extraction of vibration features and the composition of the input data for the automatic diagnosis system. Section 5 is reserved for the design of the diagnosis system, while the discussion of the results is listed in Section 6. A conclusion is presented at the end of this article.

2. Apparatus and experimentation

2.1. *Test bench*

The validation of our study was based on the exploitation of experimental data from the Bearing Data Center of Case Western Reserve University (CWRU). Figure 1 illustrates the experimental setup. The vibration signals of the bearing were collected for three of the three elements of the bearing, namely; the inner ring (IR), the rolling element (Ball) and the outer ring (OR) (see Footnote 1), this dataset is a widely recognized benchmark in bearing fault detection research. It has served as the experimental basis for hundreds of scientific publications, owing to the diversity of simulated faults and the high quality of the acquired measurements. Indeed [30], through a detailed analysis using classical signal processing methods, demonstrated that these data are easily diagnosable, making them particularly relevant for comparative evaluation of various diagnostic and classification methods and algorithms (see Footnote 1).

The test bench used for data collection illustrated in Figure 1 includes: a 2-horsepower motor (on the left), a torque sensor/encoder (in the center), a dynamometer (on the right), an electronic control box (not shown).The test bearings are mounted on the motor shaft. Localized faults were introduced into the bearings using electrical discharge machining (EDM) with diameters of 7, 14, 21, 28, and 40 mils (1 mil = 0.001 inch). SKF bearings (Table 1) were used for faults of 7, 14, and 21 mils, while NTN bearings were used for the 28 and 40 mils faults. The defective bearings were then reinstalled on the test bench, and vibration signals were recorded under various load conditions ranging from 1 to 3 horsepower, corresponding to rotational speeds between 1796 and 1720 rpm. Vibration signals were acquired at a sampling frequency of 12.000 Hz on the fan end and 48.000 Hz on the drive end. Measurements include both healthy and defective bearings exhibiting different types of faults (Figure 2). For more technical details, readers are referred to the official documentation (see Footnote 1).

Table 1. Geometric characteristics of the SKF 6205-2RS JEM bearing

Types	Values
Rolling element number	9
Ball diameter (d)	312.6 mm
Pitch diameter (D)	1537 mm
Contact angle (θ)	0

2.2. *Data used*

The primary purpose of our contribution is practical; our objective is to design a system compatible with the main ISO standards used in vibration monitoring. Specifically, we are addressing fault classification and alert thresholds for predictive maintenance of rotating machinery, according to the ISO series, notably ISO 13370 (for alert thresholds and severity assessment) and

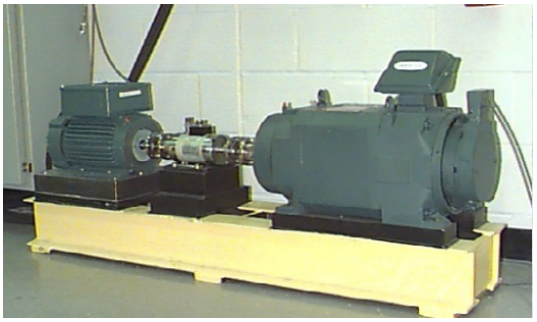


Figure 1. Experimental setup of the rolling bearings for Case Western Reserve University (see Footnote 1).

Table 2. Alert thresholds and severity assessment

Zone	Machine health status	Presence and evolution or severity of the defect	Assigned class
A	The machine is generally considered new and in good condition	No defects (healthy)	1
B	Defect zone acceptable for long-term operation	Defect 1	2
C	Defect zone not suitable for continuous long-term operation. Corrective measures must be implemented	Defect 2	3
D	Vibration levels are considered severe enough to cause damage to the machine. Operation must be stopped as soon as possible	Defect 3	4

the ISO 13374 series (for the processing and interpretation of diagnostic data). These standards define four (4) severity zones (A, B, C, D) as defined by Table 2. However, for a more rigorous approach to condition-based maintenance, ISO 10816 only provides for three classes; this case is also covered by our study. Indeed, the bearing vibration signals are annotated according to the type of defect (inner ring, outer ring, ball) and its severity. Four operating states are defined: a normal state and three levels of degradation corresponding to localized defects of increasing size (0.18 mm, 0.36 mm, and 0.53 mm). This structure allowed for the creation of three distinct datasets, each dedicated to a defective component. This approach facilitates the development and evaluation of specialized diagnostic models for each bearing element.

The states corresponding to defects of 0.71 mm and 1.02 mm were intentionally excluded because they were produced on a different bearing type (NTN), while the first three defects were machined on SKF bearings to eliminate the potential influence of the bearing type (technology, materials) on the quality of the vibration signal. This ensures a more reliable comparison between the classes and also complies with the aforementioned ISO standards. The analysis focuses exclusively on signals sampled at 12 kHz, covering all types of bearing elements (balls, outer ring, inner ring), as shown in the table below.

All signals used in this study were recorded under steady-state conditions at a constant speed of 1797 rpm (approximately 29.95 Hz). Several signal segments of equal duration were extracted for each class to create a balanced and homogeneous dataset for analysis and classification.

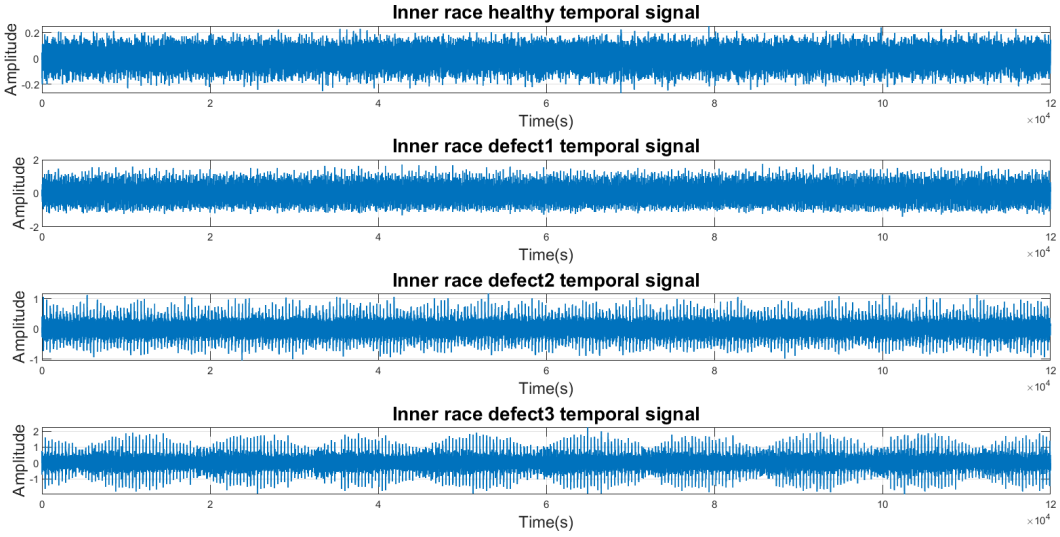


Figure 2. Temporal signal of the ball in the 4 states (healthy, defect 1, defect 2 defect 3).

Table 3. Constitution of data sets

Class	Bearing condition	Fault location		
		Outer race	Inner race	Rolling element (Ball)
1	Healthy	"OR" set	"IR" set	"Ball" set
2	Defect 1 (0.007″)	"OR" set	"IR" set	"Ball" set
3	Defect 2 (0.014″)	"OR" set	"IR" set	"Ball" set
4	Defect 3 (0.021″)	"OR" set	"IR" set	"Ball" set

3. Features extraction and analysis

3.1. *Signal preprocessing*

Before training the classification models, rigorous data preprocessing is essential to ensure the quality of the results and reduce biases related to scale or descriptor redundancy. The raw vibration signals from the database (see Footnote 1) are first preprocessed by dividing them into segments of equivalent duration to ensure representativeness and homogeneity of the samples for each class studied. All signals considered are sampled at 12 kHz. Each segment constitutes a unique observation, associated with a class corresponding to the bearing condition (healthy or with a given defect) (Table 3).

3.2. *Features extraction*

To discriminately model vibration signals from bearings, 46 features were extracted, divided into three domains: time domain, frequency domain, and time–frequency domain. The following equations provide a formal understanding of each indicator.

3.2.1. *Temporal features*

An initial signal exploration phase was conducted to gain qualitative insights into signal behavior across different defect classes. Time-domain plots of representative signals were examined to identify visible trends and irregularities. A first phase of signal exploration was carried out

in order to obtain qualitative information on its behavior according to different classes of defects. The time traces of representative signals (Figure 3) were examined in order to identify visible trends and irregularities. In order to obtain measurable indications of this information, we opted to calculate the statistical indicators below. A discrete signal is considered $x = \{x_1, x_2, \ldots, x_N\}$, where: N is the total number of samples, x_i is the signal value at the time, μ is the mean of x and σ is the standard deviation.

The statistical indicators considered relevant to describe the qualitative information contained in the signal x_i are [31,32]:

$$\text{Mean:} \quad \mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

$$\text{Variance:} \quad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2 \tag{2}$$

$$\text{Standard deviation:} \quad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2} \tag{3}$$

$$\text{RMS (Root Mean Square):} \quad \text{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i} \tag{4}$$

$$\text{Crest Factor:} \quad \text{Cr} = \frac{\max(|x_i|)}{\text{rms}} \tag{5}$$

$$\text{Skewness (asymmetry):} \quad \text{Skewness} = \frac{\sum_{i=1}^{N} (x_i - \mu)^3}{\sigma^4} \tag{6}$$

$$\text{Kurtosis (flattening):} \quad \text{Kurtosis} = \frac{\sum_{i=1}^{N} (x_i - \mu)^4}{\sigma^4} \tag{7}$$

Moments of order 4 and 5

$$m_4 = \frac{1}{N}\sum_{i=1}^{N} x_i^4 \tag{8}$$

$$m_5 = \frac{1}{N}\sum_{i=1}^{N} x_i^5 \tag{9}$$

Therefore, in the time domain, we calculated nine (9) statistical indicators for the three (03) sets of data of the constituent elements of the bearings (inner race, outer race and rolling element (ball)).

3.2.2. *Frequency features*

Subsequently, a frequency-domain analysis was performed using standard techniques such as Fast Fourier Transform (FFT), envelope spectrum analysis, and Power Spectral Density (PSD). These methods revealed distinct frequency components associated with bearing faults. Fault characteristic frequencies namely, Ball Pass Frequency of Inner Race (BPFI), Ball Pass Frequency of Outer Race (BPFO) and Ball Spin Frequency (BSF) were calculated based on the mechanical configuration of the test bench and overlaid on the spectral plots for validation [33]. These preliminary results reaffirm the effectiveness and accessibility of classical signal processing methods for fault diagnosis, in line with prior studies such as [29]. The frequency domain, obtained via the Fourier transform, provides valuable information on the energy distribution in the spectrum. Five characteristics are extracted: the dominant frequency, the average frequency, the bandwidth, the spectral energy, and the spectral kurtosis. The latter, in particular, is sensitive to abnormal energy peaks in the spectrum, often associated with localized defects in the bearings.

Table 4. Bearing defect frequencies under the rotating speed of 1796 rpm

Signal	Types	Values
Healthy signal	Shaft rotation frequency (f_r)	29.93 Hz
Signal 0.021″	Inner defect frequency (f_{id})	151 Hz
Signal 0.021″	Outer defect frequency (f_{od})	119 Hz
Signal 0.014″	Ball defect frequency (f_{bd})	126 Hz

The discrete Fourier transform (DFT) of the signal x gives the spectrum $\chi(f_k)$ for $k = 1, 2, \ldots, N_f$, where f_k is the kth discrete frequency and $S(f_k)$ is the power spectral density defined by [10]:

$$S(f_k) = |\chi(f_k)|^2 \tag{10}$$

$$\chi(f_k) = \sum_{n=0}^{N-1} x(n) \mathrm{e}^{-\mathrm{j}2\pi \frac{kn}{N}} \tag{11}$$

The features are [29]:

$$\text{Dominant frequency:} \quad f_{\text{dom}} = \text{argmax}(f_k) \tag{12}$$

$$\text{Average frequency:} \quad f_{\text{avr}} = \frac{\sum_{k=1}^{Nf} fkS(fk)}{\sum_{k=1}^{Nf} S(fk)} \tag{13}$$

$$\text{Bandwidth:} \quad \text{BW} = \frac{\sum_{k=1}^{Nf} (fk - f\text{avr})^2 S(fk)}{\sum_{k=1}^{Nf} S(fk)} \tag{14}$$

$$\text{Spectral energy:} \quad E = \sum_{k=1}^{Nf} S(fk) \tag{15}$$

$$\text{Spectral Kurtosis:} \quad \text{Kurtf} = \frac{1}{N} \sum_{k=1}^{Nf} \left(\frac{s(fk) - \mu f}{\sigma f} \right)^4 \tag{16}$$

μf and σf are the spectral mean and standard deviation in the frequency domain, we extract five (5) important characteristics which are; the dominant frequency, the average frequency, the width of the frequency bandwidth, the spectral energy and the spectral kurtosis.

To calculate the fault frequencies of a bearing, formulas must be used that depend on the geometry of the bearing and its rotational speed. These frequencies, called characteristic fault frequencies (Table 4), are: the outer ring frequency (BPFO), the inner ring frequency (BPFI) and the rolling element frequency (BSF) according to the following formula as [2]:

$$\text{Inner race defect frequency:} \quad f_{\text{id}} = \frac{N \cdot f_r}{2} \left(1 + \frac{d}{D} \cos(\theta) \right) \tag{17}$$

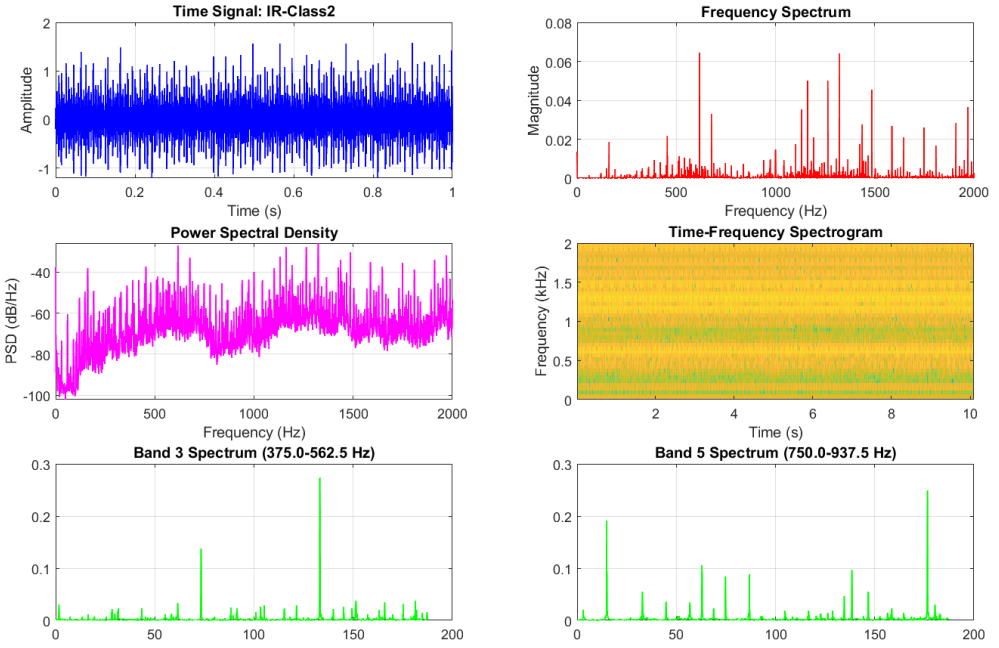
$$\text{Ball defect frequency:} \quad f_{\text{bd}} = f_r \frac{D}{d} \left(1 - \left(\frac{d}{D} \cos(\theta) \right)^2 \right) \tag{18}$$

$$\text{Outer race defect frequency:} \quad f_{\text{od}} = \frac{N \cdot f_r}{2} \left(1 - \frac{d}{D} \cos(\theta) \right) \tag{19}$$

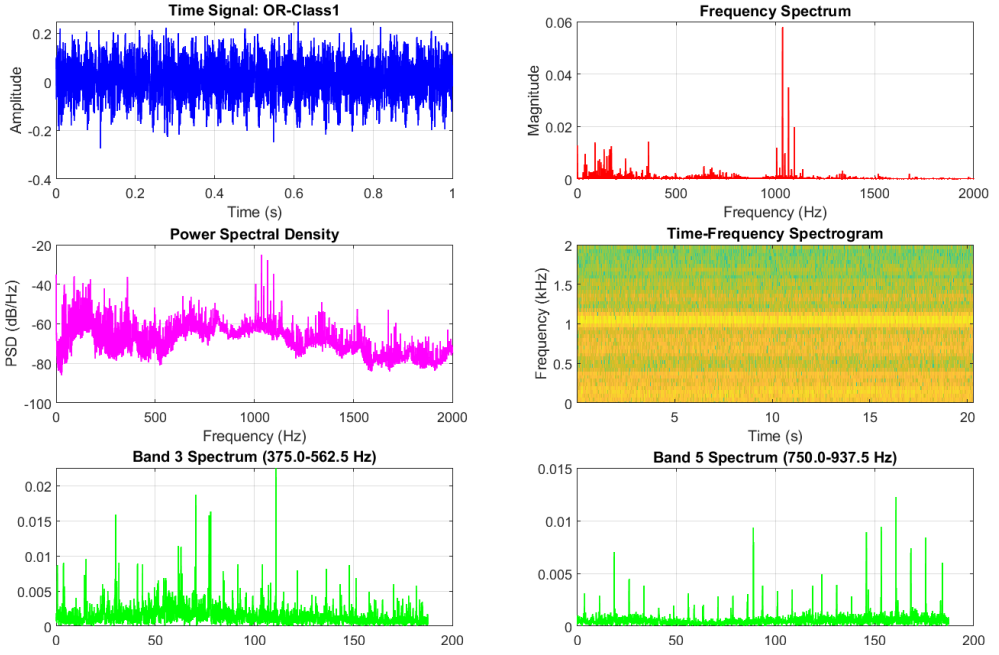
where f_r is the shaft rotation frequency, d is the ball diameter, D is the pitch diameter, N is the number of rolling elements and θ is the contact angle. Applying these formulas to our case, we found the values below.

3.2.3. *Time–frequency features*

In order to capture the temporal evolution of the signal's frequency components, we used the Wavelet Packet Transform (WPT) to obtain an accurate time–frequency representation of our

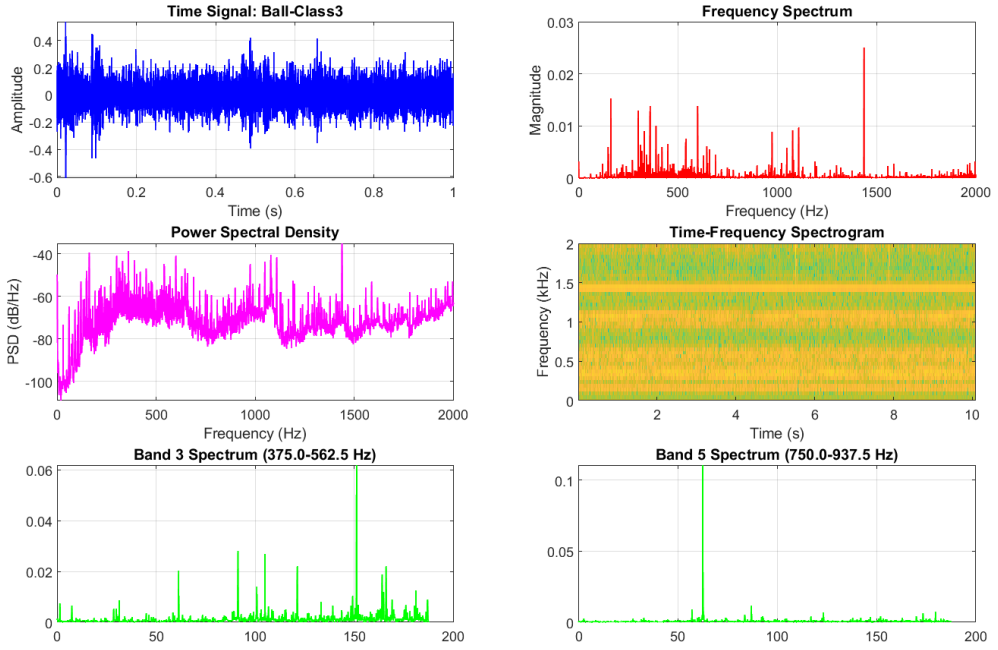


(a) IR signal



(b) OR signal

Figure 3. Continued on next page.



(c) Ball signal

Figure 3 (cont.). Time signals and their spectra.

signals. This method offers the advantage of providing an optimal joint resolution in time and frequency. The choice fell on a Daubechies wavelet of order 5 (db5), offering a good compromise between regularity and temporal support. The decomposition was carried out up to level 5, thus generating 32 terminal nodes (2^5), each corresponding to a specific frequency sub-band. To extract meaningful information and reduce the dimensionality of the data, we calculated the energy associated with each terminal node. The energy E_i of node i is defined by: energy of node [2].

$$E_i = \sum_{i=1}^{N} |x_i(k)|^2, \quad k = 1,2,\ldots,32 \tag{20}$$

where E_i is the energy of node i and $x_i(k)$ represents the kth wavelet coefficient of this node. This approach allows to obtain a vector of 32 energy characteristics, revealing the spectral distribution of the signal and its evolution over time. These energies reflect the time–frequency distribution of the signal, useful for detecting transient faults.

Finally, we obtained 32 characteristics representing the energies in nodes, which gives us thirty-two (32) time–frequency characteristics.

The time–frequency analysis was completed by examining spectrograms and kurtograms. The spectrogram (obtained by short-term Fourier transform, STFT) allows visualization of the evolution of spectral power over time, providing a general map of signal activity. The kurtogram, for its part, is a powerful tool for optimally locating transient and impulsive frequency bands, characteristic of mechanical defects. The joint analysis of these two representations aims to detect the presence of anomalies, identify the frequency band carrying the defect information and, ultimately, reveal the characteristic frequencies associated with the damage (Figures 7–11).

Table 5. Data table

Signal label	Signal class	Statistical features	Frequency features	Time–frequency features	Rpm
X97_ DE_time	1	1, 2, . . . , 10	1, 2, . . . , 5	1, 2, . . . , 32	1796
X97_ FE_time	1	1, 2, . . . , 10	1, 2, . . . , 5	1, 2, . . . , 32	1796
. . .	. . .	. . .	. . .	. . .	. . .
X105_BA_time	2	1, 2, . . . , 10	1, 2, . . . , 5	1, 2, . . . , 32	1797
. . .	. . .				

3.3. *Features engineering*

The feature engineering phase produced a total of 46 descriptors per signal sample: 9-time parameters, 5 frequency parameters and 32 energies resulting from the wavelet packet decomposition. These features were structured in a matrix of dimensions $N \times M$, where N is the number of observations and $M = 46$ the number of variables (Table 5). Each observation is also associated with a label of; class (the type of defect); signal indication and rotation speed (Table 5). This matrix is used as an input dataset for the classification algorithm. We propose for this task a novel hybrid model, merging Hidden Markov Models (HMM) with a Radial Basis Function Neural Network (RBFNN). The validation of this approach was carried out on three distinct datasets, each corresponding to a specific type of defect: outer race (OR_Race), inner race (Inner_Race) and rolling element (Ball).

3.4. *Data analysis and preparation*

3.4.1. *Class separability—Fisher score*

To assess the ability of extracted features to discriminate between different classes, we use the Fisher score, a classic measure of inter-class separability. The Fisher score is particularly relevant in supervised classification because it highlights the proportion of inter-class variance relative to intra-class variance. For a given feature x_i, the Fisher score is defined as follows [34,35]:

$$F(x_i) = \frac{\sum_{c=1}^{C} N_c(\mu_c - \mu)^2}{\sum_{c=1}^{C} \sum_{j=1}^{N_c} (\mu_{c,j} - \mu_c)^2} \tag{21}$$

where C is the number of classes, N_c the number of samples in class c, μ_c the mean of the feature in class c and μ the overall mean.

This score is calculated for each extracted feature. A high Fisher score indicates better separability of the class concerned by the feature studied. Thus, the features with the highest Fisher scores are the most discriminating and therefore the most useful for supervised learning. These scores can also guide dimensionality reduction. Indeed, this procedure allowed us to select ten (10) features with the highest Fisher scores.

3.4.2. *Dimensionality reduction*

After the initial selection of the ten (10) most relevant features via the Fisher score, a Principal Component Analysis (PCA) was deployed to address potential multicollinearity issues and achieve a more aggressive dimensionality reduction. PCA projects the data into a new orthogonal space of maximum variance, producing decorrelated principal components (PCs) [15]. Only the first k components, explaining a predefined threshold of the total variance (97%), were retained (Figure 4). This step simplifies the model structure, reduces computational time, and mitigates overfitting by eliminating noise and redundancy.

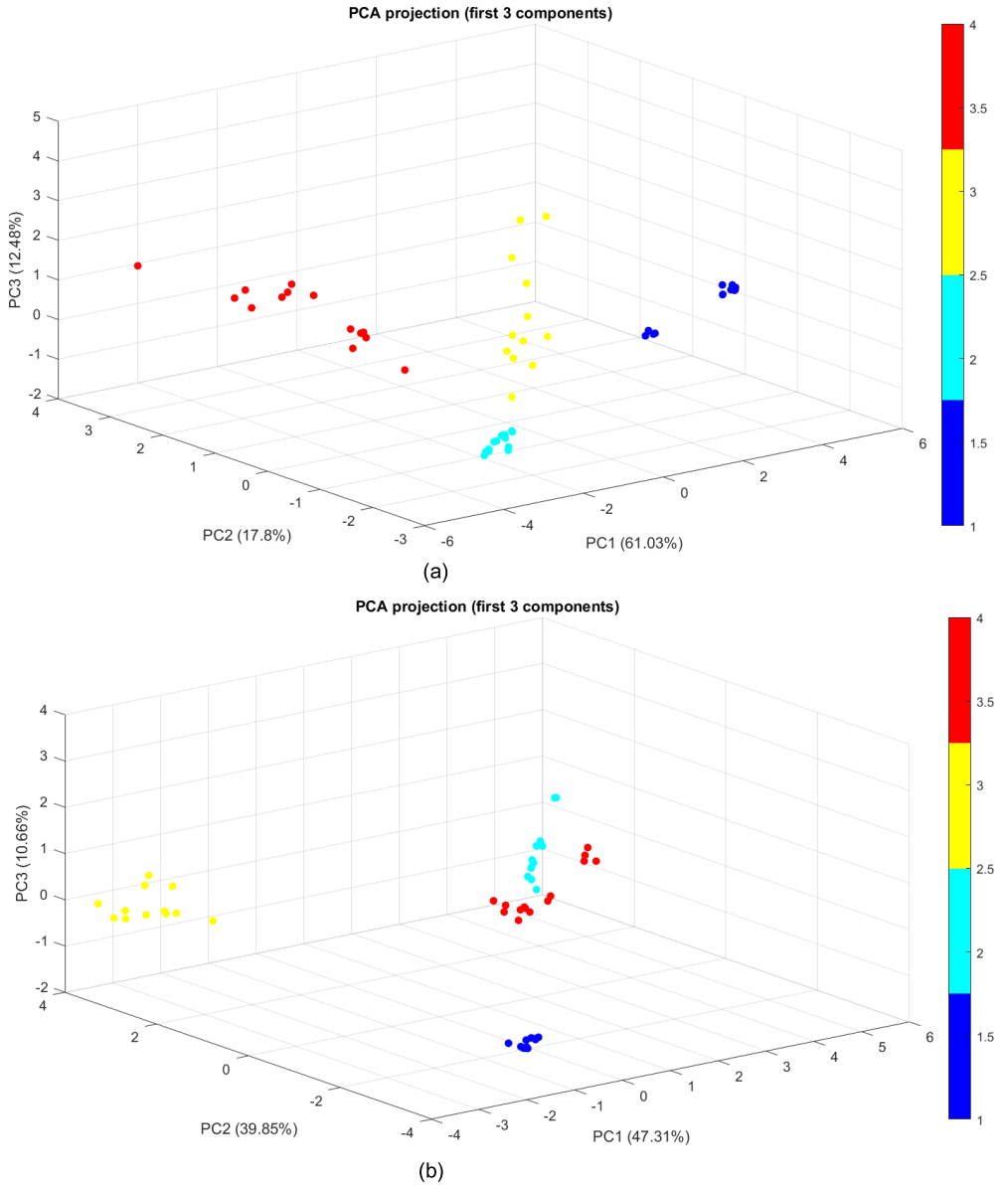


Figure 4. Continued on next page.

4. Theoretical overview of HMMs and RBFNNs

4.1. *Concept about HMM*

The Hidden Markov Model was first described in the 1960s and first used in speech recognition in the 1970s. In the late 1980s, it was applied to DNA analysis and became an important technology in the field of biological information. Through continued exploration and application of this technology, it now finds widespread applications in many fields such as fault diagnosis, machine learning, automated driving, natural language processing, and target recognition. The HMM is a statistical model for describing a hidden-state Markov process. It must first be trained based on existing data [29].

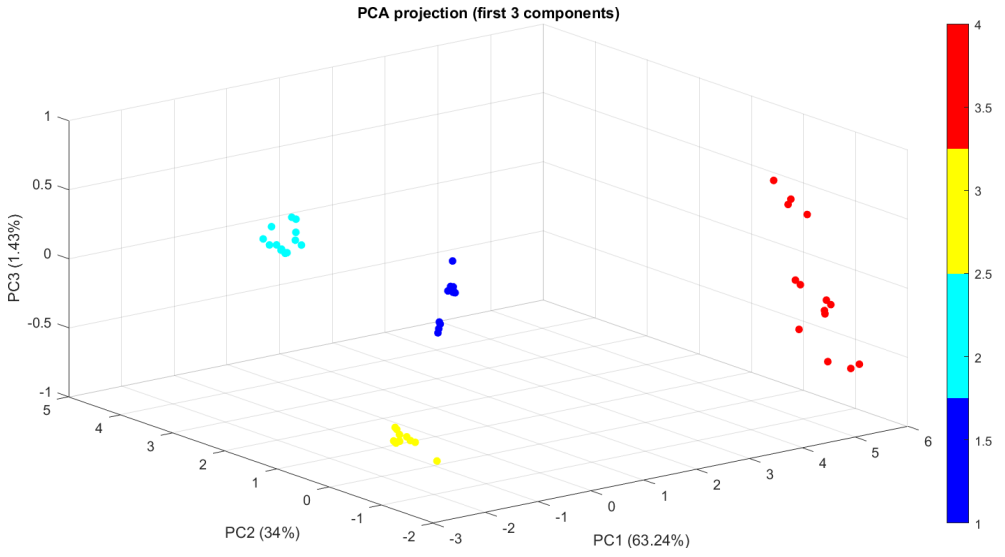


Figure 4 (cont.). Dimensionality reduction by PCA (a) Ball features, (b) Inner_race features and (c) Outer_race features.

An HMM is a doubly stochastic process with an underlying Markov process that is not observable (the states are hidden), but can only be observed through another set of stochastic processes that are produced by the Markov process (the observations are probabilistic functions of states). The observation values are used to determine the model parameters [3]. These values are then analyzed and identified according to the established model.

HMMs can be defined by the following parameters: number of hidden states N, number of observable states M corresponding to each hidden state, initial probability distribution π, state transfer probability matrix A and emission probability matrix B or probability density function [6,22], so we can therefore define an HMM as: $\lambda = (A, B, \pi)$. An HMM can also be represented graphically by different structures, including the example in Figure 1 representing the structure used in our case, called the left-right model or Bakis model. This choice models the irreversible phenomenon of wear on the bearing elements. Indeed, wear progresses forward without reversing, with probabilities estimated by the Markov process described below; therefore, the probabilities of backward wear are practically zero. The Markov states S1, S2, S3, and S4 in Figure 5 represent the classes that can potentially affect the bearing elements, which is a hidden process, while the outputs Y1, Y2, Y3, and Y4 represent the bearing's health states, informing us of the bearing's current state, which is a visible process based on probabilities estimated by the Markov process.

4.1.1. *Markov hypothesis*

The future state of a process does not depend on the past state, but only on the present state [3,22,29].

$$q_{i+1} = f(q_i), \quad i = 1, \ldots, t \tag{22}$$

4.1.2. *Immobility hypothesis*

The output of the system is not related to time [3,29].

$$P(q_{j+1}|q_j) = P(q_{i+1}|q_i), \quad i \neq j, j \in 1, \ldots, t \tag{23}$$

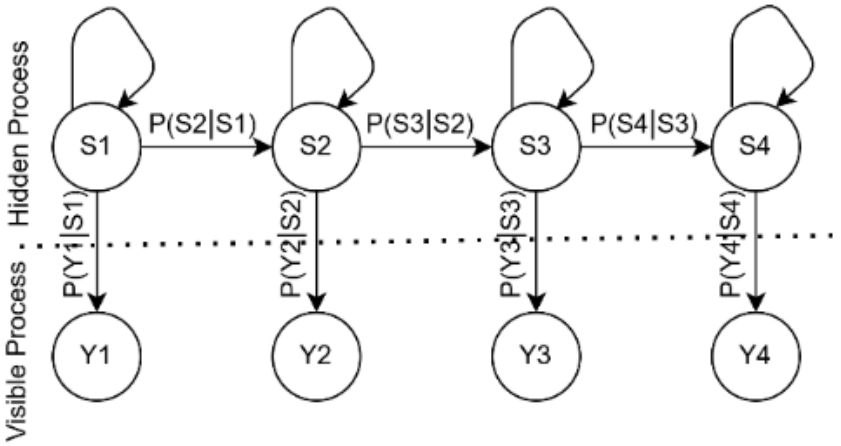


Figure 5. Structure of the 4-state Markov model (left-right).

4.1.3. *Output independence hypothesis*

The output of the system is only related to the current state of the system [3,29].

$$P(o_1, o_2, \ldots, o_t | q_1, q_2, \ldots, q_t) = \prod_{i=1}^{t} P(o_i | q_i) \tag{24}$$

In this paper, based on the estimation of model parameters (Table 6) and the determination of the location and severity of fault signals by maximum likelihood, the forward algorithm, the inverse algorithm, and the forward-inverse algorithm are adopted. The definition and symbols of each algorithm are as follows [3,29].

4.1.4. *Forward algorithm*

The forward variable is defined as the probability of reaching a given state, given the first t observations of the sequence [22,29]:

$$\alpha(t, i) = P(o_1, o_2, \ldots, o_t, Q_t = q_i | \lambda), \quad 1 \le t \le T \tag{25}$$

where T is the length of observations. The forward recursion is expressed as [1,22]

$$\alpha(t+1) = \left[\sum_{i=1}^{N} \alpha_j(t) a_{ij} \right] B_j(o_{t+1}) \tag{26}$$

4.1.5. *Backward algorithm*

The backward variable is defined as the probability of observing the remaining observations given any starting point t [36,37]

Backward algorithm: the backward variable is defined as the probability of observing the remaining observations given a starting point t.

$$\beta(t, i) = P(o_{t-1}, o_{t-2}, \ldots, o_T, Q_t = q_i | \lambda), \quad 1 \le t \le T-1 \tag{27}$$

4.1.6. *The backward recursion is expressed as [1,22]*

$$\beta(t) = \left[\sum_{i=1}^{N} a_{ij} B_j(o_{t+1}) \right] B_j(t+1) \tag{28}$$

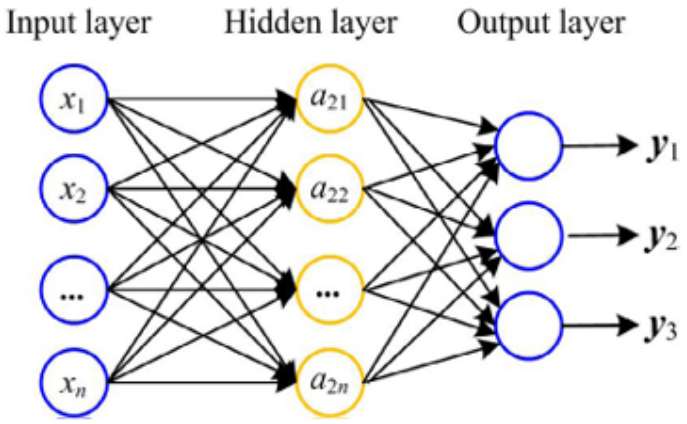


Figure 6. Structure of an RBFNN [25].

4.1.7. *Forward–backward algorithm (Baum–Welch algorithm)*

Forward–backward algorithm obtains a set of forward probabilities and a set of backward probabilities, wich can be used to jointly acquire the distribution over states at any specific time t [1,22]:

$$\begin{aligned}\xi_t(i,j) = P(Q_t = q_i, Q_{t+1} = q_j | O, \lambda) &= \frac{\alpha_t(i) A_{ij} B_j(o_{t+1}) \beta_{t+1}(j)}{P(O|\lambda)} \\ &= \frac{\alpha_t(i) A_{ij} B_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i) A_{ij} B_j(o_{t+1}) \beta_{t+1}(j)} \\ & 1 \le i, j \le N, j \le t \le T \end{aligned} \tag{29}$$

HMMs mainly solve three application problems: evaluation, decoding and learning. The first is to calculate the log-likelihood (LL) when the parameters and observation sequences are known. The second is to find the most probable state sequence when the observation sequence is known. The third is to adjust the parameters to maximize the emission probability of the observation sequence. The classic methods to solve these three problems are the forward–backward algorithm, the Viterbi algorithm and the Baum–Welch algorithm [35].

4.2. *Concept about RBFNN*

Radial basis function (RBF) networks are similar to multilayer perceptron (MLP) networks. They also have direct unidirectional connections, and each neuron is fully connected to the units in the next layer. The neurons are organized in a layered topology with direct dynamics (from inputs to outputs, without loops). RBF networks differ fundamentally in how they model the relationship between inputs and outputs. While MLP networks model this relationship in a single step, an RBF network divides this learning process into two independent and distinct steps (Figure 6). First, using the neurons in the hidden layer called radial basis functions, RBF networks model the probability distribution of the input data. Second, the RBF network learns to relate the input data x to a target variable t. Note: unlike MLP networks, the bias term in an RBF neural network is connected only to the output neurons. In other words, RBF networks do not have a bias term linking the inputs to the radial basis units [25].

Table 6. Simulation of classes using HMM parameters

Class designation	HMM designation	Maximum likelihood
Class 1	λ_1	Log $(P(O \mid \lambda_1))$
Class 2	λ_2	Log $(P(O \mid \lambda_2))$
Class 3	λ_3	Log $(P(O \mid \lambda_3))$
Class 4	λ_4	Log $(P(O \mid \lambda_4))$

An RBF is represented by the structure in Figure 2; the outputs are defined by Equation (9) [25].

$$y_k = \sum_{j=1}^{M} \omega_j(k) \cdot \varphi_j(x) + b \tag{30}$$

$$\varphi_j(x) = \mathrm{e}^{-\frac{(x-\mu_j)^2}{\sigma_j^2}} \tag{31}$$

where $\omega_j(k)$ are the weights optimized by gradient descent, b is the bias.

Learning in an RBFNN consists of:

- Choose the centers μ_j (for example k-means),
- Determine the widths σ_j,
- Estimate the weights $\omega_j(k)$ (by linear or pseudo-inverse regression)

Radial basis networks can be used to approximate functions. newrb adds neurons to the hidden layer of a radial basis network until it reaches the specified mean squared error target.

5. Supervised classification methodology for bearing defects

The overall classification methodology is summarized in the flowchart in Figure 7. This entire process is independently applied to each of the three datasets corresponding to localized faults on the outer race, the inner race, and the rolling element (ball). This independent treatment allows the capture of specific structural characteristics related to each type of mechanical defect, ensuring a more accurate domain-specific modeling.

6. Bearing defect classification

6.1. *Bearing Defect Classification Using HMMs*

Supervised defect classification is orchestrated by a battery of hidden Markov models (HMMs), following a so-called "one-HMM-per-class" architecture. For each health state (healthy, 0.007″ defect, 0.014″ defect, 0.021″ defect) of each component (OR, IR, Ball), a separate HMM, denoted λ_k, is fully learned from its training data. The fundamental principle of this learning is maximum log-likelihood estimation. The objective is to adjust the model parameters $\lambda = (A, B, \pi)$, respectively the state transition matrix, the emission (or observation) matrix and the initial distribution, in such a way that it becomes the most likely generator of the observation sequences $O = (o_1, o_2, o_3, o_4)$ presented to it.

Formally, this amounts to solving the following optimization problem for each class:

$$\tilde{\lambda} = \mathrm{argmax}(\log(P(O \mid \lambda_i))), \quad i = 1, 2, 3, 4 \tag{32}$$

In practice, the probability $P(O \mid \lambda)$ itself is extremely complex and difficult to maximize directly. This is why the Baum–Welch algorithm, a specific instance of the EM (Expectation–Maximization) algorithm, is employed. This iterative algorithm actually maximizes the logarithm

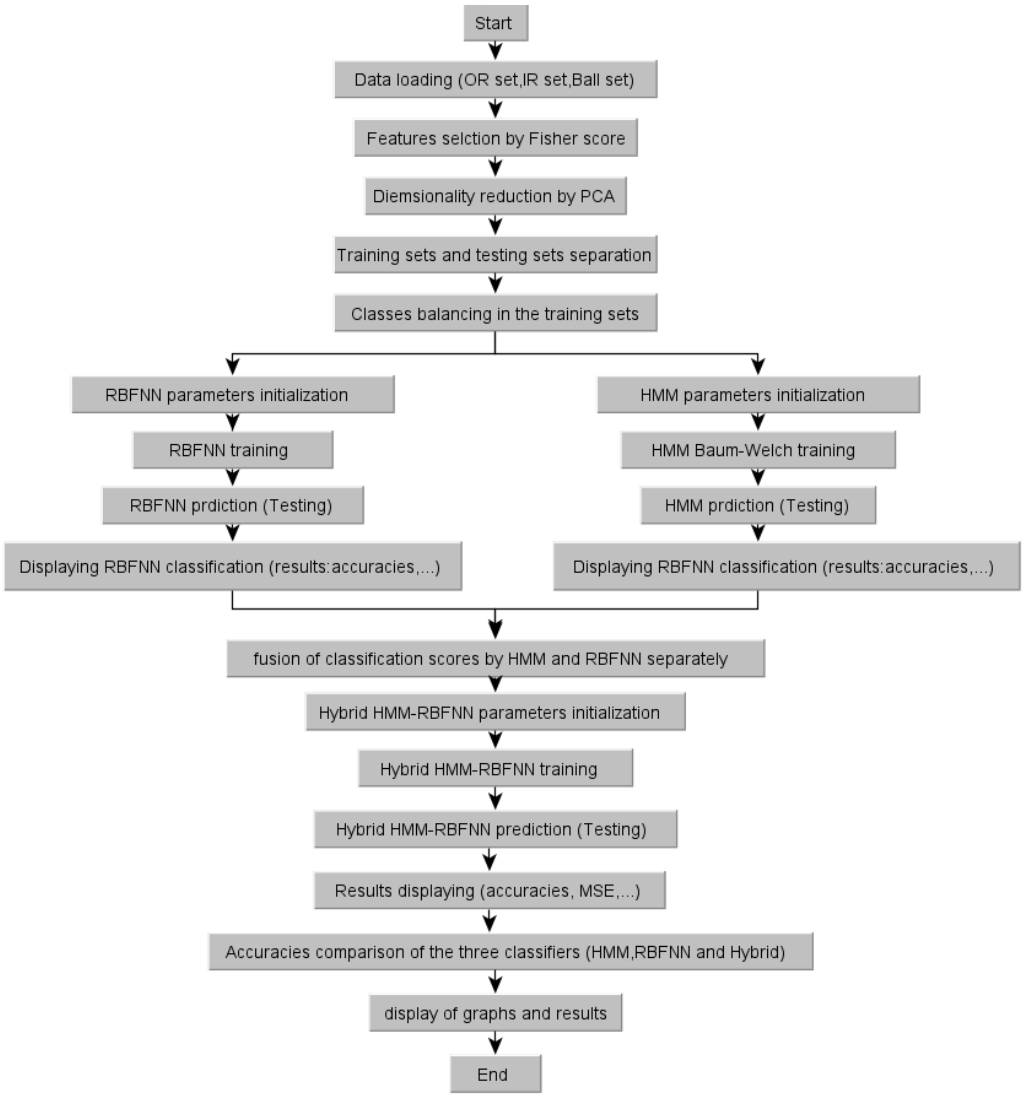


Figure 7. Implementation flowchart of the hybrid HMM-RBFNN bearing defect classification system.

of the likelihood, according to relation (32), which is mathematically equivalent but numerically more stable, avoiding the overhead of product probability calculations.

The E (Expectation) process calculates, a current estimate of the parameters λ, the forward (α) and backward (β) probabilities in order to estimate the expected value of the hidden state assignment and transitions. The M (Maximization) process then uses these expectations to update the parameters $\lambda(A,B,\pi)$ in order to maximize this expected likelihood. These two steps are repeated until convergence, typically when the increase in the log-likelihood $\mathscr{L}(\lambda)$ between two iterations falls below a predefined threshold.

The concrete significance of this process for our application is profound; for example, by maximizing $P(O \mid \lambda_1)$ for the healthy model, we "teach" the HMM the normal dynamic and probabilistic signature of a bearing. Conversely, by maximizing $P(O \mid \lambda_2)$ for defect model 1, we encap-

sulate in its parameters (A, B) the unique "stochastic vibration signature" of an incipient 0.007-inch defect, i.e., the probable frequency sequences and their characteristic temporal concatenation. Thus, each HMM becomes a compact and probabilistic representation of the vibration personality of the class it represents. In the testing phase, classifying an unknown signal with its sequence of observations "O" simply amounts to calculating the log-likelihood $\log(P(O \mid \lambda_k))$ for each HMM trained using the forward algorithm. The class $\tilde{k}$ that maximizes this measure, $\tilde{k} = \text{argmax}(\log(P(O \mid \lambda_k)))$, is assigned to the signal, because its model is the one that best explains, from a probabilistic point of view, the genesis of the observed sequence.

6.2. *Bearing defect classification using Radial Basis Function Neural Networks (RBFNN)*

6.2.1. *Label preprocessing*

To prepare the data for supervised learning, the discrete class labels are first transformed into a format suitable for neural networks using one-hot encoding. This technique represents each class as a binary vector where a single element is active (value 1), indicating class membership, while the others are inactive (value 0). This transformation is crucial because it allows the RBFNN to output a vector of scores or probabilities for each class, rather than a simple discrete label.

6.2.2. *Training the RBF architecture*

The RBFNN is then trained on the previously normalized training set. The learning process relies on adjusting several critical hyperparameters:

- The dispersion parameter (spread or sigma) of the radial basis function (often a Gaussian), which controls the width and influence of each hidden neuron.
- The number of neurons in the hidden layer, which determines the model's complexity.
- The error tolerance, which defines the algorithm's convergence criterion.

A common approach is to build the network incrementally: neurons are iteratively added to the hidden layer until the reconstruction error on the training data falls below the defined tolerance threshold, thus optimizing the model's ability to generalize.

6.2.3. *Prediction phase*

During the testing phase, each sample is presented to the network. The output layer produces a vector of scores (or activations) for all classes. The decision rule adopted is simple: the predicted class is the one associated with the highest score (max rule). The performance of the RBFNN classifier is quantified using standard metrics as illustrated in the following section. This evaluation protocol, identical to that used for Hidden Markov Models (HMM), allows an objective and rigorous comparison of the performances between the three approaches; HMM, RBFNN and hybrid HMM-RBFNN.

6.3. *Bearing defect classification using the hybrid HMM-RBFNN system*

This hybrid system was designed to leverage the complementary strengths of Hidden Markov Models (HMMs) excellent for modeling temporal dependencies in sequences and RBF Neural Networks—powerful for capturing complex nonlinear relationships in point-like data. The central idea is to merge their opinions to make a more robust and accurate decision.

6.3.1. *Score fusion*

For each test sample, the two classifiers independently produce their own scores.

• HMM Side

The model generates a log-likelihood score for each class. Since these log-likelihood scores are not normalized and can vary considerably in scale, they are transformed into a probability distribution using the Softmax function. This produces a probability vector P_HMM of size 4, where each element is between 0 and 1 and the sum of the elements is 1.

• RBFNN side

The neural network directly produces a score vector (or sometimes probabilities if an activation function like Softmax is used as output). This vector is denoted S_RBF for the 4 classes.

The fusion is performed by a weighted average of these two score vectors, after scaling them (usually $[0,1]$).

• Fusion formula for class k

$$\text{Score}_{\text{fusion}}(k) = \alpha \cdot P_{HMM}(k) + (1-\alpha) \cdot S_{\text{RBFNN}}(k), \quad \alpha \epsilon [0,1] \tag{33}$$

Parameter α controls the relative influence of each model on the final decision. Several α values, both below and above 0.5, were experimentally evaluated to analyze their impact on classification performance. The results show that the value ($\alpha = 0.5$) offers the best compromise between the temporal modeling provided by the HMM and the nonlinear discrimination capability of the RBFNN, thus leading to the best overall classification performance. This observation confirms the value of a balanced fusion that complementarily leverages the strengths of both approaches.

Hybrid model training and prediction. The approach described above is a decision-level fusion (or "score fusion").

Creating the fusion inputs. A new dataset is created. For each training example, the HMM (P_HMM) and RBFNN (S_RBF) score vectors are concatenated to form a new feature vector of length 8 (4 + 4).

Retraining. This new "fusion" feature vector becomes the input for a final classifier. Often, a simple model such as logistic regression or RBFNN is used. This hybrid classifier is retrained using the original one-hot labels. Its role is to learn how and when to trust each model to make the best collective decision.

Prediction. To leverage the complementary capabilities of the HMM and RBFNN models, a decision fusion hybridization scheme was implemented. The outputs of the two classifiers are merged using a weighting coefficient α, according to relation (33).

6.4. *Results and discussion*

This section presents a detailed comparative analysis (Table 7) of the performance of the three classification architectures namely; Hidden Markov Models (HMM), Radial Function Neural Networks (RBFNN) and the hybrid HMM-RBFNN system. For the classification of bearing defects, the evaluation is conducted separately for the three critical elements: Ball, Inner Ring (IR) and Outer Ring (OR), aiming to discriminate the four health states (Class 1 : Healthy, Class 2: Defect 1, Class 3: Defect 2, Class 4: Defect 3).

Table 7. Comparison of overall accuracy (%)

Classifier	Accuracy (%)		
	OR	IR	Ball
HMM	91.00	98.66	95.33
RBFNN	95.00	100.00	98.66
Hybrid HMM-RBFNN	99.00	100.00	100.00

Table 8. Detailed performance (Accuracy per class) of HMM, RBFNN and Hybrid classifiers for each element of the bearing

Class	HMM			RBFNN			Hybrid HMM-RBFNN		
	OR	IR	Ball	OR	IR	Ball	OR	IR	Ball
Class 1 (healthy)	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Class 2 (defect 1)	80.00%	100.00%	100.00%	92.00%	100.00%	100.00%	96.00%	100.00%	100.00%
Class 3 (defect 2)	100.00%	100.00%	94.70%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Class 4 (defect 3)	84.00%	94.70%	86.80%	88.00	100.00%	94.70%	100.00%	100.00%	100.00%

6.5. *Overall performance*

Table 8 summarizes the overall accuracy of each feature and each classifier. A clear and consistent trend emerges from these results.

The hybrid system achieves perfect or near-perfect performance (100% for Ball and IR, 99% for OR), consistently outperforming both models used alone. This superiority unequivocally demonstrates the effective synergy between the HMM, which excels at modeling the temporal evolution of vibration signatures, and the RBFNN, powerful at capturing complex nonlinear relationships in the feature space. Their complementary strengths allow the hybrid system to make a more robust and accurate decision, correcting the errors that each model makes individually.

6.6. *Performance analysis by class and interpretation*

A granular analysis by class (Table 8), reveals the nature of the errors and the strengths of each approach.

6.7. *Consistent performance of the healthy state*

Regardless of the classifier or the tested element, Class 1 (Healthy) is always identified with 100% accuracy. This crucial result indicates that none of the models produces a false positive (a healthy part identified as defective). This characteristic is essential for an industrial diagnostic system, as it avoids unnecessary and costly maintenance downtime.

6.8. *Systematic superiority of the hybrid system*

The hybrid model made no errors on the inner ring (IR) and ball (Ball) test sets, achieving 100% accuracy across all their defect classes. For the outer ring (OR) (Figure 8), it significantly improved

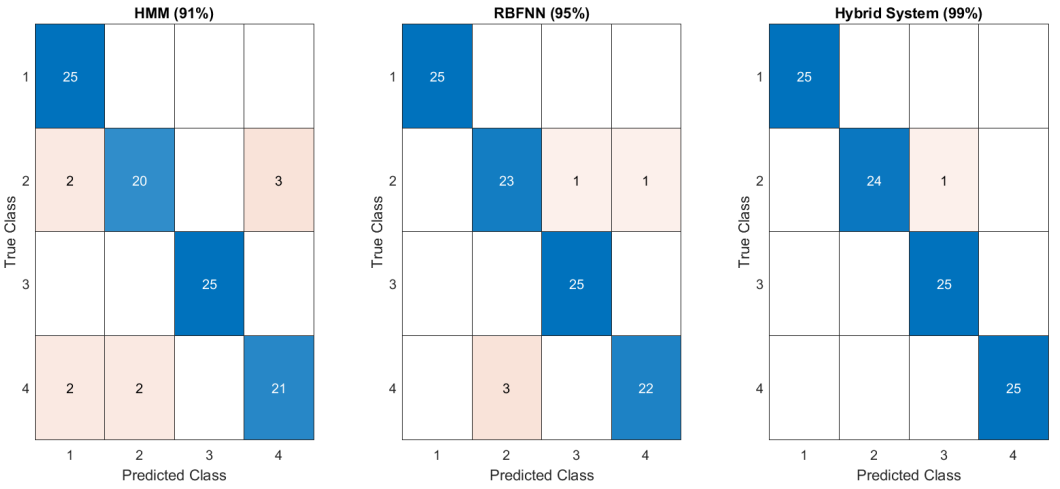


Figure 8. OR Confusion matrices.

the results of the individual models, increasing the accuracy to 96% for Defect 1 and 100% for Defect 3. This perfect or near-perfect performance demonstrates that merging the models corrects their respective errors and creates a more robust and reliable system.

6.9. *Comparative analysis of individual models*

• RBFNN vs HMM

The RBFNN outperforms the HMM in 100% of cases where the HMM does not reach 100%. For example, for Defect 3 on the OR, the RBFNN (88%) outperforms the HMM (84%) by 4 points. This trend confirms the greater nonlinear modeling power of the RBFNN to capture complex relationships in the data, whereas the sequence-based HMM shows its limitations.

• Identified weaknesses of the HMM

The weaknesses of the HMM focus on specific defects. OR Defect 1 (80%): The HMM's lower performance suggests that the vibration signature of this particular defect may be less periodic or noisier, making it difficult to model with a Markov chain.

Ball (86.8%) and OR (84%) Defect 3: These results confirm that the HMM struggles with certain severe defects, likely because they generate complex signals that encroach on the "space" of other classes in the sequential model.

• The Outer Ring (OR): the most critical element to diagnose

The lowest results, across all classifiers (before hybridization), are observed for the OR (Defaults 1 and 3). This identifies the OR as the most challenging element to model and classify among the three. Hybridization provides the most dramatic added value here, fully addressing the deficiencies of the individual models.

6.10. *Performance conclusion*

This analysis unequivocally demonstrates that while the individual models (HMM and RBFNN) are already highly efficient, the hybrid HMM-RBFNN system is the architecture of choice for

Table 9. Maximum log-likelihood of observation sequences for rolling element and class-trained HMM models

Class	Maximum Log-likelihood Log ($P(O \mid \lambda_i)$)		
	Outer race	Inner race	Ball
Class 1 (healthy)	−53129.73	−3967.20	−80048.80
Class 2 (defect 1)	−1154.20	−353.19	−946.80
Class 3 (defect 2)	−94.23	−290.34	−2031.18
Class 4 (defect 3)	−1131.56	−358.32	−205.50

reliable and accurate bearing fault diagnosis. It optimally combines the strengths of both approaches:

- The HMM's ability to model temporal dynamics.
- The RBFNN's ability to learn complex nonlinear decision boundaries.
- The robustness of the fused system, which eliminates single points of failure.

The hybrid system guarantees near-infallible fault detection and identification, regardless of the faulty component, making it a cutting-edge solution for predictive maintenance.

Table 9 presents the maximum log-likelihood ($\text{Log}(P(O \mid \lambda_i))$) of the test data for the HMM models specific to each class and rolling element.

6.11. *Differential performance by element*

Modeling complexity varies considerably depending on the rolling element. Models for the inner ring (IR) exhibit overall less negative and more homogeneous log-likelihoods (from −3967 to −290) than those for the outer ring (OR) and ball (Ball), suggesting that the vibration signatures of IR defects are more distinctive and easier to model by HMMs. This corroborates the overall high performance of the classifiers on the IR.

6.12. *Healthy state modeling*

For all three elements, the "Healthy" class (Class 1) consistently exhibits the lowest (most negative) log-likelihood. This indicates that the vibration signal of a healthy bearing is more complex, noisier, and exhibits greater variability, making it inherently more difficult to model with a high probability of an HMM compared to the periodic and structured signals generated by localized defects.

6.13. *Identification of the most distinctive defects*

For each element, one defect stands out with significantly better modeling performance:
OR—Class 3 (Defect 2): The exceptionally high value (−94.23) indicates that this defect generates an extremely stereotypical vibration signature that is perfectly captured by the HMM. This pattern is so specific that it risks "attracting" sequences belonging to other classes, explaining possible confusion. IR—Class 3 (Defect 2): Shows the best log-likelihood (−290.34), confirming that it is the most characteristic defect for this element. Ball—Class 4 (Defect 3): Has the highest log-likelihood (−205.50), which corresponds to excellent detection of the most severe defects on the ball.

6.14. *Confusion point prediction*

Defect pairs with close log-likelihoods are potential sources of confusion for the HMM classifier alone. For example: For OR: Classes 2 and 4 have very similar values (−1154.20 vs. −1131.56), which suggests a risk of confusion between Defect 1 and Defect 3. For Ball: A significant gap exists between Class 4 (−205.50) and Class 3 (−2031.18), suggesting that Defect 2 is much more difficult to correctly identify than Defect 3, which is reflected in the class-specific accuracy results.

6.15. *Conclusion*

This log-likelihood analysis provides a fundamental explanation for the performance of the HMM classifier. It highlights the most distinctive defects, the most difficult classes to model (healthy), and predicts classification ambiguities. It thus fully justifies the need to use a non-linear classifier like the RBFNN, and ultimately the hybrid system, to compensate for these limitations and achieve optimal and robust classification accuracy.

7. Multi-base generalization

Having demonstrated the effectiveness of the HMM-RBFNN hybrid system on the CWRU database, widely used as a reference in the literature for bearing fault diagnosis, it becomes necessary to evaluate the generalizability of the proposed approach. Indeed, a high-performing classification method must not only deliver good results on a given database but also maintain its performance when applied to datasets from different test benches, operating conditions, and sensors. With this in mind, the study is extended to the Paderborn and Axial Bearing databases, which present experimental characteristics and fault scenarios significantly different from those of the CWRU database.

The Paderborn and Axial Bearing databases were integrated to evaluate the robustness and generalizability of the HMM-RBFNN hybrid system. Unlike the CWRU database, which relies primarily on artificially localized defects and relatively controlled operating conditions, the Paderborn database includes real defects resulting from progressive damage processes, as well as greater variability in rotational speeds and mechanical loads. The Axial Bearing database, on the other hand, is distinguished by its analysis of vibration signals measured in the axial direction, thus providing a complementary evaluation framework based on a different mechanical configuration and vibration dynamics. Applying the HMM-RBFNN hybrid system to these two databases therefore allows us to verify its ability to effectively exploit features extracted from signals exhibiting different statistical distributions, frequency contents, and temporal dynamics.

7.1. *Paderborn University (PU) database*

The Paderborn database[2], developed by the University of Paderborn (Germany), is a reference database for studying bearing defect diagnosis under conditions closely resembling real-world industrial environments. Unlike many databases based on artificially created defects, this database includes real defects resulting from progressive damage processes. Measurements are performed on a modular test bench (Figure 9) incorporating an electric motor, a belt drive system, a bearing-supported shaft, and a controlled loading device. Vibration signals (Figure 10) are collected using accelerometers mounted on the bearing housing at various rotational speeds and

[2]https://mb.uni-paderborn.de/en/kat/research/bearing-datacenter.

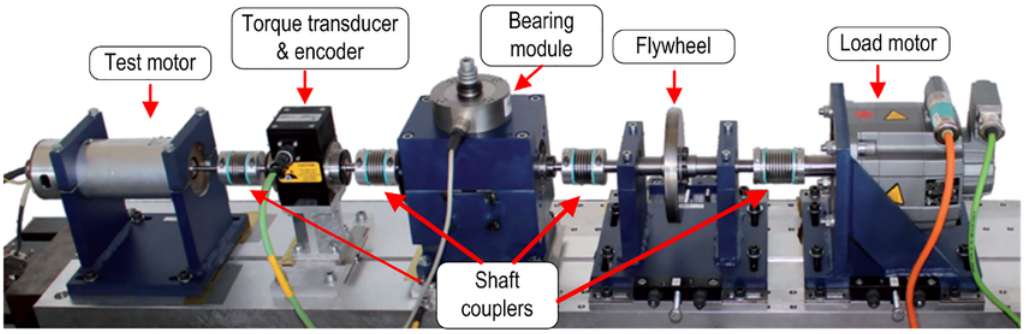


Figure 9. Paderborn test bench (see Footnote 2).

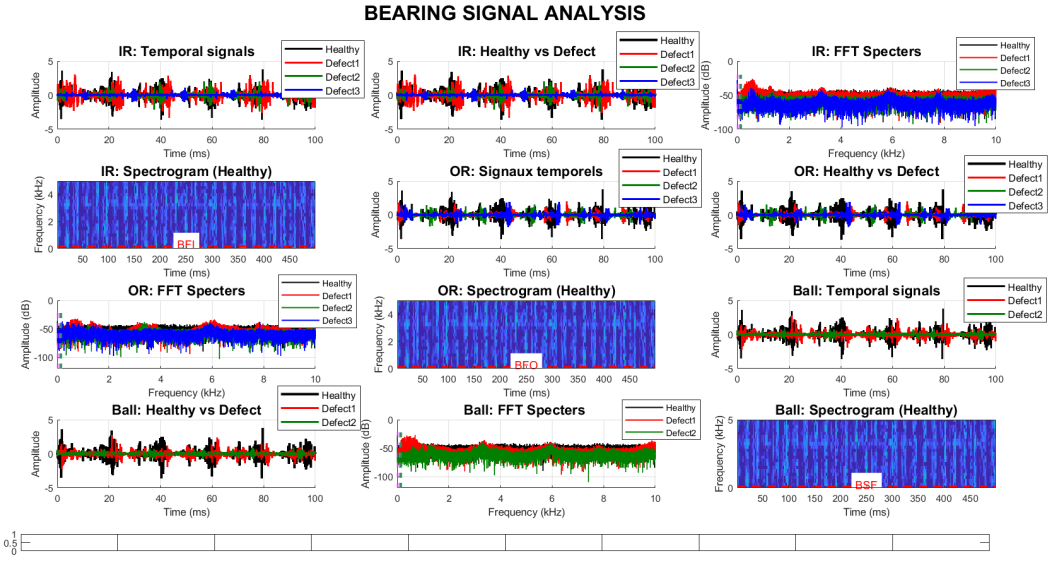


Figure 10. Paderborn data base signal analysis.

load levels. The classes studied include sound bearings as well as several types of localized defects, including inner and outer race defects.

Following the same procedure established for the CWRU database, the same analysis process was carried out to arrive at the results below.

7.1.1. *Fault classification*

Table 10 presents the classification performance by class and the overall accuracy obtained by three approaches: HMM, RBFNN, and the hybrid HMM-RBFNN system. The results show that the HMM model achieves perfect accuracy (100%) for all classes, confirming its ability to effectively model the temporal dynamics of vibration signals. The RBFNN model also exhibits excellent overall performance (99.50%), although a slight degradation is observed for class 2 (97.50%), indicating increased sensitivity to overlapping features between closely related fault states.

Table 10. Classification accuracy (%)

	HMM	RBFNN	Hybrid HMM-RBFNN
Class 1	100.00	100.00	100.00
Clas 2	100.00	97.50	100.00
Class 3	100.00	100.00	100.00
Overall	100.00	99.50	100.00

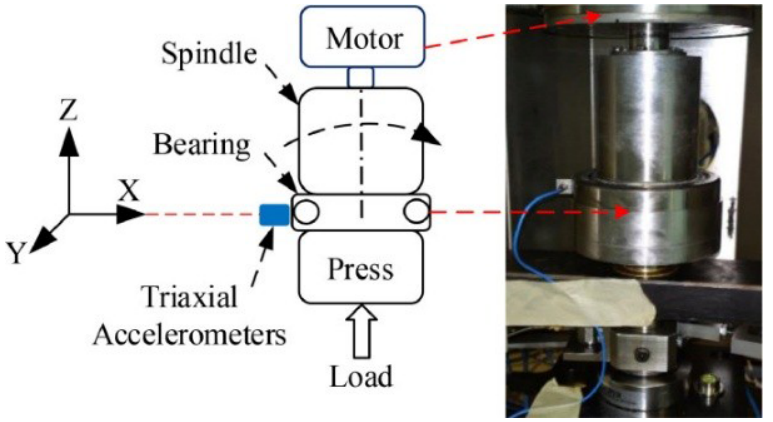


Figure 11. Schematic and photograph of a special bearing testing machine [38].

7.1.2. *Conclusion*

These results demonstrate that the HMM-RBFNN hybridization surpasses individual approaches by guaranteeing perfectly stable classification for all classes. The combined integration of temporal information and nonlinear generalization capabilities thus constitutes an effective and reliable strategy for bearing fault diagnosis, reinforcing the relevance of the hybrid system for real-world industrial applications. In contrast, the hybrid HMM-RBFNN system restores 100% accuracy for all classes by combining the sequential modeling offered by HMM with the strong nonlinear discrimination capabilities of RBFNN. This complementarity improves the robustness of the diagnosis, particularly for classes that are more difficult to separate.

7.2. *Axial bearing database*

The Axial Bearing database [38], is dedicated to the analysis of bearing defects subjected to axial loads, thus offering a complementary framework to traditional databases primarily focused on radial vibrations. The test bench (Figure 11) consists of an electric motor driving a shaft supported by ball bearings, on which various axial load conditions are applied. Spall-type defects are introduced on the bearing elements to simulate realistic degradation scenarios.

Vibration measurements (Figure 12) are performed using accelerometers positioned in the axial direction, enabling the capture of specific vibration signatures often overlooked in conventional diagnostic approaches.

The results obtained (Table 11) show that the hybrid approach maintains high performance across all the databases studied, thus confirming its robustness in the face of variations in experimental conditions and data structures.

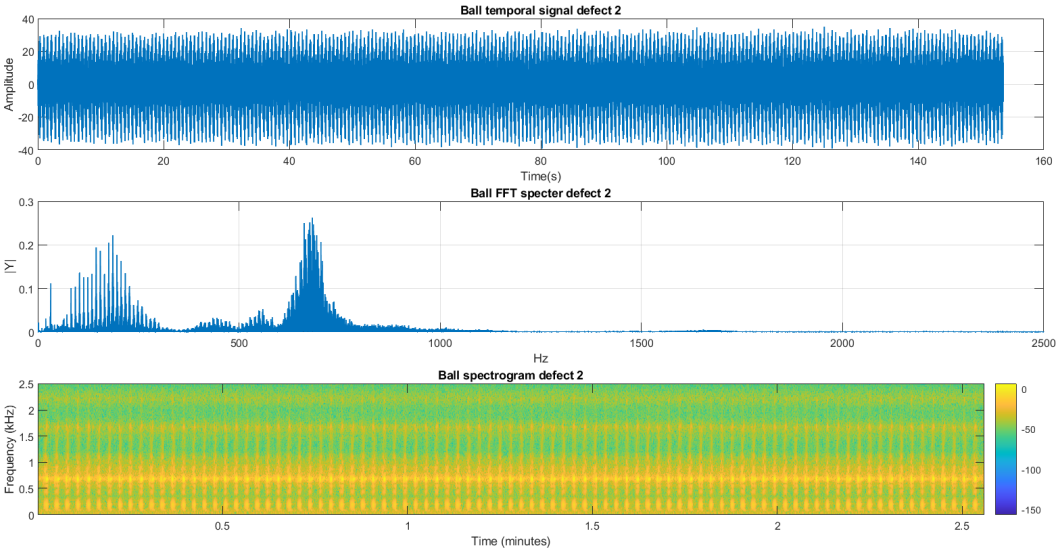


Figure 12. Axial bearing data base Ball signal analysis.

Table 11. Classification accuracy (%)

Classifier	OR	IR	Ball
HMM	78.667	78.667	78.667
RBFNN	100	100	100
Hybrid	100	100	100

Table 12. Performance of the HMM-RBFNN hybrid system as a percentage (%) for the three databases studied

Classifier	Database		
	CWRU	Paderborn	Axial bearing
HMM	94.99	98.65	78.667
RBFNN	97.88	98.78	100
Hybrid	99.66	99.33	100

The results of the Table 12 confirm that the HMM-RBFNN hybrid system is not limited to performance specific to a given database, but constitutes a generalizable and robust diagnostic approach, capable of adapting to varied experimental contexts and different bearing configurations

7.2.1. *Recapitulation*

The summary Table 12 presents the classification performance obtained by the three approaches studied (HMM, RBFNN, and Hybrid HMM-RBFNN) on three reference datasets: CWRU, Paderborn, and Axial Bearing.

The results show that HMM offers good performance on the CWRU (94.99%) and Paderborn (98.65%) datasets, confirming its ability to model the temporal dynamics of vibration signals. However, its accuracy drops significantly in the Axial Bearing dataset (78.67%), indicating increased sensitivity to the variability of experimental conditions and the complexity of axial defects.

RBFNN improves overall performance on all three datasets, particularly on Axial Bearing where it achieves perfect accuracy (100%), thanks to its strong nonlinear feature separation capability. Nevertheless, its performance remains slightly lower than that of the hybrid system on the CWRU and Paderborn datasets.

The hybrid HMM-RBFNN system consistently achieves the best or equivalent maximum performance across all databases, with classification rates of 99.66%, 99.33%, and 100%, respectively. These results confirm the robustness and generalizability of the hybrid model, resulting from the complementarity between the temporal modeling of HMM and the discriminatory power of RBFNN.

7.3. *Comparison with the state of the art*

The performance of the proposed hybrid system is compared with recent work in the field. The results obtained (Table 13) are competitive with the most recent and efficient approaches published in the literature. The system achieves some accuracy equivalent to that of hybrid models based on deep convolutional neural networks (CNN), which are known to be very data- and computationally intensive. This fully validates the choice of the HMM-RBFNN hybridization as an efficient, robust and elegant architecture for the bearing fault diagnosis task.

Table 13. Comparison with hybrid methods in the literature

Reference	Hybrid method	Acronym decoding/description	Reported accuracy (%)
[27]	Hybrid FMM-RF	Fuzzy Min-Max (FMM) neural network and the Random Forest (RF) model,	99.89
[37]	Hybrid CNN-MLP model	Convolutional Neural Network-Multi Layer perceptron	98.00
[39]	Hybrid CNN-MLP	Convolutional Neural Network-Multi Layer perceptron	100.00
[40]	PCA + Hybrid NN	PCA: Principal Component Analysis (dimensionality reduction); NN: Neural Network	≈98.70
[41]	CWT-CNN-BiLSTM-Attention	CWT: Continuous Wavelet Transform; BiLSTM: Bidirectional Long Short-Term Memory; Attention: feature weighting	>99.00
[42]	IHHO-DBN-ELM	IHHO: Improved Harris Hawks Optimization; DBN: Deep Belief Network; ELM: Extreme Learning Machine	≈99.50
[43]	CNN-Dual Feature Selection + SHAP	SHAP: SHapley Additive explanations (model interpretability)	>99.00
Our contribution	HMM-RBFNN	Hidden Markov Models-Radial Basis Function Neural Network	≈99.67
[44]	Deep CNN + Random Forest	Deep CNN: Convolutional Neural Network (automatic feature extraction); Random Forest: ensemble learning classifier	≈99.20

Note: A weighted average of the results on Ball, IR and OR gives an overall accuracy of 99.69%.

7.4. *Summary of findings*

Merging the scores of the HMM and RBFNN models proved to be a winning strategy. It compensates for the individual weaknesses of each model and leverages their complementary strengths to produce a more robust and accurate classifier, capable of diagnosing any type of fault on the various bearing components with an exceptionally high success rate.

8. General conclusion

This comparative study confirms that hybrid diagnostic frameworks significantly improve bearing fault classification by combining complementary modeling capabilities. Deep learning-based hybrids, particularly CNN-based architectures, exhibit remarkable accuracy but often require large training datasets and substantial computing resources. Conversely, the proposed HMM-RBFNN approach offers competitive performance while maintaining reduced complexity and better adaptability to data-constrained environments. By leveraging the time-domain modeling power of Hidden Markov Models and the nonlinear classification capability of radial basis function neural networks, the hybrid system effectively captures the dynamic and discriminating characteristics of vibration signals. The consistent performance achieved on the CWRU, Paderborn, and Axial Bearing datasets underscores the robustness and generalizability of the proposed method. These results demonstrate that the HMM-RBFNN hybrid is a reliable and efficient alternative to complex deep learning solutions for practical bearing fault diagnostic applications.

Future work and perspectives

While this work opens promising avenues, several areas for improvement and research can be considered:

- Automatic hyperparameter optimization: Further exploration, via Bayesian optimization algorithms or expanded search grids, could further refine performance, particularly by optimizing the fusion weighting factor α.
- Generalization of more varied data: It would be relevant to test the model's robustness on datasets from different machines, operating under varied conditions (load, speed, noise levels) to assess its generalization capacity.
- Exploration of other fusion architectures: Testing earlier fusion, at the feature level rather than at the score level, could potentially lead to the discovery of even more discriminating patterns.
- Towards an embedded system: The next step would be to optimize the final model for deployment on an embedded platform, enabling real-time diagnostics directly on site, which would represent a significant advancement for industrial predictive maintenance.

In conclusion, this work not only enabled the development of a powerful classifier; it also highlighted the immense potential of hybrid architectures to address the complex challenges of Industry 4.0, combining the strengths of multiple models to create a solution that is smarter and more reliable than the sum of its parts.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

References

[1] G. Xin, N. Hamzaoui and J. Antoni, "Semi-automated diagnosis of bearing faults based on a hidden Markov model of the vibration signals", *Measurement* **127** (2018), pp. 141–166.

[2] P. K. Kankar, S. C. Sharma and S. P. Harsha, "Fault diagnosis of ball bearings using continuous wavelet transform", *Appl. Soft Comput.* **11** (2011), no. 2, pp. 2300–2312.

[3] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", *Proc. IEEE* **77** (1989), no. 2, pp. 257–286.

[4] Y. Chen, T. Zhang, Z. Luo and K. Sun, "A novel rolling bearing fault diagnosis and severity analysis method", *Appl. Sci.* **9** (2019), no. 11, article no. 2356.

[5] B. Samanta, K. Al-Balushi and S. Al-Araimi, "Artificial neural networks and genetic algorithm for bearing fault detection", *Soft. Comput.* **10** (2006), pp. 264–271.

[6] J. Wang, G. Du, Z. Zhu, C. Shen and Q. He, "Fault diagnosis of rotating machines based on the EMD manifold", *Mech. Syst. Signal Process.* **135** (2020), article no. 106443.

[7] J. B. Ali, B. Chebel-Morello, L. Saidi, S. Malinowski and F. Fnaiech, "Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network", *Mech. Syst. Signal Process.* **56–57** (2015), pp. 150–172.

[8] S. Tang, S. Yuan and Y. Zhu, "Deep learning-based intelligent fault diagnosis methods toward rotating machinery", *IEEE Access* **8** (2020), pp. 9335–9346.

[9] H. A. Gabbar, "Design of intelligent fault diagnostic system (FDS)", *Process Saf. Environ. Prot.* **85** (2007), no. 6, pp. 566–578.

[10] Z. Huo, Y. Zhang, P. Francq, L. Shu and J. Huang, "Incipient fault diagnosis of roller bearing using optimized wavelet transform based multi-speed vibration signatures", *IEEE Access* **5** (2017), pp. 19442–19456.

[11] Q. Jiang, Y. Shen, H. Li and F. Xu, "New fault recognition method for rotary machinery based on information entropy and a probabilistic neural network", *Sensors* **18** (2018), article no. 337.

[12] P. Jayaswal, S. N. Verma and A. K. et Wadhwani, "Development of EBP-Artificial neural network expert system for rolling element bearing fault diagnosis", *J. Vib. Control* **17** (2011), no. 8, pp. 1131–1148.

[13] D. Liu, W. Cheng and W. Wen, "An online bearing fault diagnosis technique via improved demodulation spectrum analysis under variable speed conditions", *IEEE Syst. J.* **14** (2020), no. 2, pp. 2323–2334.

[14] J. Sun, C. Yan and J. Wen, "Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning", *IEEE Trans. Instrum. Meas.* **67** (2018), no. 1, pp. 185–195.

[15] Y. Gu, X. Q. Zhou and D. P. Yu, "Fault diagnosis method of rolling bearing using principal component analysis and support vector machine", *J. Mech. Sci. Technol.* **32** (2018), pp. 5079–5088.

[16] B. Sreejith, A. K. Verma and A. Srividya, "Fault diagnosis of rolling element bearing using time-domain features and neural networks", in *2008 IEEE Region 10 and the Third International Conference on Industrial and Information Systems, Kharagpur, India*, 2008, pp. 1–6.

[17] T. Xie, H. Yu and B. Wilamowski, "Comparison between traditional neural networks and radial basis function networks", in *2011 IEEE International Symposium on Industrial Electronics, Gdansk, Poland*, 2011, pp. 1194–1199.

[18] J. Fan, L. Wei, W. Zhongqiu, W. Zewen and C. Baoyu, "Fault diagnosis of rotating machinery based on MFES and D–S evidence theory", in *Proceedings of the 24th Chinese Control and Decision Conference (CCDC), Taiyuan*, 2012, pp. 1624–1629.

[19] Y. Wu, H. Wang, B. Zhang and K.-L. Du, "Using radial basis function networks for function approximation and classification", *ISRN Appl. Math.* **2012** (2012), article no. 324194.

[20] G. Georgoulas, T. Loutas, C. D. Stylios and V. Kostopoulos, "Bearing fault detection based on hybrid ensemble detector and empirical mode decomposition", *Mech. Syst. Signal Process.* **41** (2013), no. 1–2, pp. 510–525.

[21] D. Kateris, D. Moshou, X.-E. Pantazi, I. Gravalos, N. Sawalhi and S. Loutridis, "A machine learning approach for the condition monitoring of rotating machinery", *J. Mech. Sci. Technol.* **28** (2014), no. 1, pp. 61–71.

[22] V. H. Pham, S. Han, M. D. Do, et al., "A wavelet packet spectral subtraction and convolutional neural network based method for diagnosis of system health", *J Mech. Sci. Technol.* **33** (2019), pp. 5683–5687.

[23] M. Sedira, R. Ziani and A. Felkaoui, "Comparison between hidden Markov models and artificial neural networks in the classification of bearing defects", in *Rotating Machinery and Signal Processing. SIGPROMD'2017* (A. Felkaoui, F. Chaari and M. Haddar, eds.), Applied Condition Monitoring, vol. 12, Springer: Cham, 2017.

[24] R. K. Jha and P. D. Swami, "Intelligent fault diagnosis of rolling bearing and gear system under fluctuating load conditions using image processing technique", *J. Mech. Sci. Technol.* **34** (2020), pp. 4107–4115.

[25] R. Liu, B. Yang, E. Zio and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: a review", *Mech. Syst. Signal Process.* **108** (2018), pp. 33–47.

[26] X. Zhang, G. Chen, T. Hao and Z. He, "Rolling bearing fault convolutional neural network diagnosis method based on casing signal", *J. Mech. Sci. Technol.* **34** (2020), no. 6, pp. 2307–2316.

[27] M. Seera, M. L. Dennis Wong and A. K. Nandi, "Classification of ball bearing faults using a hybrid intelligent model", *Appl. Soft Comput.* **57** (2017), pp. 427–435.

[28] J. Liu, Y. Hu, B. Wu, Y. Wang and F. Xie, "A hybrid generalized hidden Markov model-based condition monitoring approach for rolling bearings", *Sensors* **17** (2017), no. 5, article no. 1143.

[29] W. Zhao, T. Shi and L. Wang, "Fault diagnosis and prognosis of bearing based on hidden Markov model with multi-features", *Appl. Math. Nonlinear Sci.* **5** (2020), no. 1, pp. 71–84. Sciendo.

[30] W. A. Smith and R. B. Randall, "Review rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study", *Mech. Syst. Signal Process.* **64–65** (2015), pp. 100–131.

[31] J. Hou, Y. Wu, A. S. Ahmad, H. Gong and L. Liu, "A novel rolling bearing fault diagnosis method based on adaptive feature selection and clustering", *IEEE Access* **9** (2021), pp. 99756–99767.

[32] V. Sharma and A. Parey, "A review of gear fault diagnosis using various condition indicators", *Procedia Eng.* **144** (2016), pp. 253–263.

[33] J. Wang, M. Xu, C. Zhang, B. Huang and F. Gu, "Online bearing clearance monitoring based on an accurate vibration analysis", *Energies* **13** (2020), article no. 389.

[34] J. Lee, F. Wu, W. Zhao, M. Ghaffari, L. Liao and D. Siegel, "Prognostics and health management design for rotary machinery systems—reviews, methodology and applications", *Mech. Syst. Signal Process.* **42** (2014), pp. 314–334.

[35] Y. Qifeng, C. Longsheng and M. T. Naeem, "Hidden Markov Models based intelligent health assessment and fault diagnosis of rolling element bearings", *PLoS One* **19** (2024), no. 2, article no. e0297513.

[36] B. Randall and J. Antoni, "Rolling element bearing diagnostics—a tutorial", *Mech. Syst. Signal Process.* **25** (2011), no. 2, pp. 485–520.

[37] V. Sinitsin, O. Ibryaeva, V. Sakovskaya and V. Eremeeva, "Intelligent bearing fault diagnosis method combining mixed input and hybrid CNN–MLP model", *Mech. Syst. Signal Process.* **180** (2022), article no. 109454.

[38] M. A. A. Ismail, J. Windelberg, A. Bierig, I. Bravo and A. Arnaiz, "Ball bearing vibration data for detecting and quantifying spall faults", *Data Brief.* **47** (2023), article no. 109019. PMID: 36942099; PMCID: PMC10023968.

[39] Y. Gao, C. H. Kim and J.-M. Kim, "A novel hybrid deep learning method for fault diagnosis of rotating machinery based on extended WDCNN and long short-term memory", *Sensors* **21** (2021), no. 19, article no. 6614.

[40] Z. You, J. Zhang and S. Li, "Fault diagnosis of rolling bearings using PCA-based hybrid neural networks", *Measurement* **189** (2022), article no. 110462.

[41] N. Siddique, M. A. Khan and J.-M. Kim, "Attention-based CWT–CNN–BiLSTM framework for intelligent bearing fault diagnosis", *Mech. Syst. Signal Process.* **195** (2025), article no. 110893.

[42] Y. Sun, H. Wang, X. Chen and L. Zhang, "An optimized hybrid IHHO–DBN–ELM model for rolling bearing fault diagnosis", *Expert Syst. Appl.* **236** (2025), article no. 121353.

[43] M. E. Saouli, M. M. Touba and A. Boudiaf, "Hybrid diagnostic framework for interpretable bearing fault classification using CNN and dual-stage feature selection", *Sensors* **25** (2025), no. 20, article no. 6386.

[44] G. Xu, M. Liu, Z. Jiang, D. Söffker and W. Shen, "Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning", *Sensors* **19** (2019), no. 5, article no. 1088.