



**HAL**  
open science

# Bearing Fault Event-Triggered Diagnosis using a Variational Mode Decomposition-based Machine Learning Approach

Houssem Habbouche, Yassine Amirat, Tarak Benkedjough, Mohamed Benbouzid

► **To cite this version:**

Houssem Habbouche, Yassine Amirat, Tarak Benkedjough, Mohamed Benbouzid. Bearing Fault Event-Triggered Diagnosis using a Variational Mode Decomposition-based Machine Learning Approach. IEEE Transactions on Energy Conversion, 2021, pp.1-9. 10.1109/TEC.2021.3085909 . hal-03249339

**HAL Id: hal-03249339**

**<https://hal.science/hal-03249339v1>**

Submitted on 1 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Bearing Fault Event-Triggered Diagnosis using a Variational Mode Decomposition-based Machine Learning Approach

Houssem Habbouche, Yassine Amirat, *Senior Member, IEEE*, Tarak Benkedjough, Mohamed Benbouzid, *Fellow, IEEE*,

**Abstract**—The monitoring of rolling element bearing is indexed as a critical task for condition-based maintenance in various industrial applications. It allows avoiding unscheduled maintenance operations while decreasing their cost. For this purpose, various methodologies were developed to ensure accurate and efficient monitoring. In this context, this paper proposes an approach for bearing fault early diagnosis based on the variational mode decomposition (VMD), used as a notch filter for dominant mode cancellation, and a machine learning approach, namely the one-dimensional convolution neural network (1D-CNN), for detection and diagnosis purposes. Specifically, the proposed approach first performs features extraction using VMD for fault detection, and then triggers to multi-scale features extraction using CNN convolution and pooling layers for classification and diagnosis.

The proposed bearing fault detection and diagnosis approach is evaluated, in terms of robustness and performances, using the well-known Case Western Reserve University experimental dataset. In addition, performances are evaluated versus well-established demodulation techniques, in terms of fault detection, and machine learning strategies, in terms of fault diagnosis. The achieved results show that the proposed VMD notch filter-based 1D-CNN approach is clearly promising for bearing degradation monitoring.

**Index Terms**—Bearing fault, convolution neural network, fault detection and diagnosis, variational mode decomposition, machine learning.

## I. INTRODUCTION

**I**NDUSTRIAL rotating machines key and critical components are clearly gears and bearings as their failure modes can lead to prolonged downtime and substantial additional maintenance costs. Monitoring the health status of these components is therefore a high industrial priority [1], [2].

Monitoring and analyzing these sensed vibrations is therefore given a great industrial importance, in addition to carry information about the rotating machine dynamic state [3]. Indeed, industrial experience feedback has shown that 70% of rotating machine shutdowns are caused by induced vibrations and 30% of them are a consequence of bearing failures [4],

[5].

Rotating machines state of health indicators have to be adequately chosen to help in the monitoring process: failure detection [6], failure diagnosis, which is in fact the classification of the failure type, and the prognosis to estimate the remaining useful lifetime in order to avoid unplanned shutdowns [7], [8]. For this purpose, many techniques and tools were developed for these failures detection and diagnosis, as well as potentially attempting to prolong the working life cycle. While the literature is rich of model-based failures detection and diagnosis approaches, data-driven or signal-based processing techniques become more attractive as they do not require any prior knowledge of rotating machines parameters in addition of their ability to extract useful features for fault recognition [9] without specific information about the rotating machine operating conditions [10]. Depending on the operating conditions, specific signal processing techniques were developed, where the most popular ones are time- and frequency-analysis techniques for steady-state conditions, while time-frequency and time-scale representations are adopted for non-stationary behaviors. As highlighted in the literature, these techniques have some limitations mainly in terms of complexity, poor resolution, and cross-terms occurrence. To address this issues, parametric methods, based on parameter estimations of a known model were developed. Nevertheless, these methods have drawbacks as they are formulated through integral transforms and analytic signal representations [11]. Hence, they depend on data length and model accuracy.

To overcome the above-mentioned drawbacks, data driven approaches based on mode decomposition (i.e. empirical mode decomposition (EMD), ensemble empirical mode decomposition EEMD, waveform mode decomposition (WPD), variational mode decomposition (VMD), etc.) were introduced and their merits highlighted as one of the most suitable bearing fault detection and diagnostic approaches given the generated non-stationary vibration signals [4], [12], [13]. These techniques are often coupled with an artificial intelligence-based algorithm for automatic fault classification [14]. Indeed, artificial intelligence, namely machine learning, is introduced to process the acquired signals and the extracted features with the objective to design a fully automated diagnosis process [13]. In this context, artificial intelligence-based techniques provide relevant tools in terms of features recognition, detection, diagnosis, and

H. Habbouche and T. Benkedjough are with Ecole Militaire Polytechnique, Mechanical Structures Laboratory, Bordj Elbahri 16046 Algiers, Algeria (e-mail: habbouche.houssem@gmail.com; bktarek@gmail.com).

Y. Amirat is with the ISEN Yncréa Ouest, L@bIsen, Brest, France (e-mail: yassine.amirat@isen-ouest.yncrea.fr).

M. Benbouzid is with the University of Brest, UMR CNRS 6027 IRDL, 29238 Brest, France, and also with the Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China (e-mail: mohamed.benbouzid@univ-brest.fr).

prediction accuracy (prognosis) [15]. Specifically, machine learning-based algorithms, such as deep neural network (DNN), recurrent neural network (RNN), and conventional neural network (CNN) are considered as the best path for features classification [16]. Indeed, these techniques efficiently extract features by useful information compression [7], [8]. Hybridizing mode decomposition with machine learning for fault detection and diagnosis, respectively, has been recently addressed in the literature. Indeed, Chen *et al.* [17] proposed a comparative study between VMD and EMD-based support vector machine (SVM) approaches for classification. They specifically investigated the entropy of each intrinsic mode function and mode as an input vector for classification purpose for wind turbine monitoring. For gearbox monitoring, Li *et al.* [18], used VMD for features extraction, by power spectral entropy, and DNN for classification. In this case, comparisons have been carried out versus back propagation neural network (BPNN) and SVM, although modes after decomposition were stationary. Gai *et al.* [19] used hybrid grey wolf optimization to find the VMD best parameters combination to improve mode decomposition for DBN-based learning and bearing faults classification. In this study, where VMD has been compared to EMD, it has been shown that convolution layers to extract more useful information can improve the diagnosis accuracy. Sharma *et al.* [20] compared VMD to empirical wavelet transform and flexible analytic wavelet transform for gearbox faults detection. This study highlighted the VMD effectiveness in clearly exhibiting a fault by extracting its transients. In the same context, Gu *et al.* [4] proposed a framework based on VMD to decompose signals, then calculating statistical indicators used as SVM inputs for learning and bearing states classification. In this case, VMD-based decomposition has been compared to EMD. However, replacing signals with temporal indicators limits the technique robustness. For the monitoring of a refrigeration system, Wang *et al.* [21] proposed a combination of 1D-CNN and gate recurrent unit (GRU) to ensure features extraction and learning features for classification. In this case, the proposal has been evaluated versus BPNN, CNN, and long short-term memory (LSTM). A multi-signal fault detection and diagnosis approach has been proposed by Hao *et al.* [22] which used 1D-CNN for features extraction that LSTM for classification. The effectiveness of this approach has been assessed versus SVM, k-nearest neighbors (KNN), BPNN, and CNN. The denoising issue has been addressed by Liu *et al.* [23] who proposed a combination of 1D-convolutional autoencoder for denoising input signals, which are then learned by 1D-CNN for faults classification. In this case, high diagnosis accuracy has been achieved. Finally, the idea to remove the signal processing step has been proposed by Jiang *et al.* [24]. It is in then replaced by convolution layers and pooling for classes discrimination. Despite machine learning obtained results for filtering and features extraction, signal processing remains very useful especially in highly noisy environment to improve diagnosis accuracy [25].

According to the above-discussed context and state of the art review, this paper proposes a new method for bearing fault

diagnosis. It is based on the VMD, as a notch filter for fault detection, and the 1D-CNN for classification purpose. In this context, the proposal main contributions are the following:

- Providing an intelligent decision support tool for real-time bearing diagnosis, ensuring detection first and fault classification later;
- High-level features extraction by filtering the dominant mode using VMD to identify the fault even in presence of high harmonics pollution;
- Multi-scale features extraction using CNN convolution and pooling layers to extract most discriminating features between different classes;
- Experimental evaluation and validation using the Case Western Reserve University (CWRU) database.

The paper is organized as follows; section II is devoted to the methodology theoretical background presentation. Section III deals with the experimental evaluation and validation. Section IV provides an analysis and discussion of the achieved results while section V concludes this paper.

## II. PROPOSED FAULT EVENT-TRIGGERED DIAGNOSIS METHODOLOGY

The proposed bearing fault diagnosis methodology is illustrated by Fig. 1 flowchart, where a VMD approach is adopted for fault detection then triggering to diagnosis using a machine learning approach, namely 1D-CNN.

The following subsections will detail the operating flow of the proposed diagnosis methodology.

### A. Variational Mode Decomposition-based Fault Detection

As above-mentioned in the state of the art review, signal processing is the step of choice to handle the issue of acquired signals corrupted by noise and harmonics. This is particularly the case of vibration signals related to mechanical components generating low amplitudes pulses [25]. Signal processing techniques are therefore used to isolate these components [2], [20]. Among the adaptive mode decomposition family, the VMD is proposed. Indeed, it has been introduced to improve the EMD and becomes the technique of choice for the analysis of non-stationary and nonlinear data for detection in a wide range of applications [26], [27], [28], [29]. VMD has the advantageous ability to decompose complex signals into several stationary signals, regardless of their origin, using Wiener filter [13].

$$x_n(t) = \sum_{i=1}^k u_k(t) + res(t), \quad (1)$$

where  $x_n$  is the acquired signal,  $\{u_k\} = \{u_1, u_2, \dots, u_n\}$  are decomposition modes, and  $res$  is the residual signal after optimisation.

The decomposition process lies in solving an optimization problem formulated as:

$$\min_{\{u_k\}, \{\omega_k\}} \left\{ \sum_k \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\}, \quad (2)$$

subject to  $\sum_k u_k = f$ .

where  $f$  is the original signal,  $\{\omega_k\}$  are center frequencies

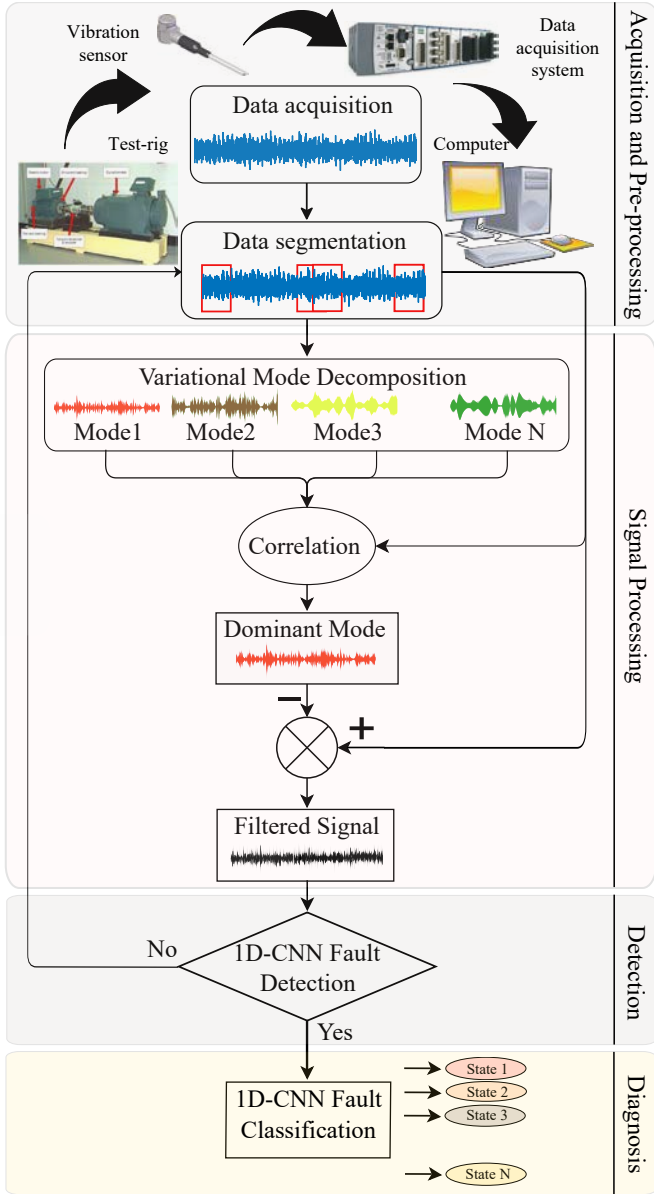


Fig. 1. Flowchart of the proposed fault event-triggered detection and diagnosis methodology.

of each  $\{u_k\}$ ,  $\delta(t)$  is an impulse function, and  $k$  is modal component number.

The new formulation of the variational constrained problem is an augmented Lagrangian equation formulated as follows [18]:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \quad (3)$$

where  $\lambda$  is the Lagrange multiplier, and  $\alpha$  is a quadratic penalty factor.

Resolution is done by iterative techniques that allow estimating modes  $u_k$  and their central frequencies  $\omega_k$  as well as the Lagrangian operator  $\lambda(t)$ , formulated iteratively in (4), (5),

and (6), respectively [30]

$$\hat{u}_k^{n+1}(\omega) \leftarrow \frac{\hat{f}(\omega) \sum_{i < k} \hat{u}_i^{n+1}(\omega) - \sum_{i > k} \hat{u}_i^n(\omega) + \frac{\lambda^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (4)$$

$$\omega_k^{n+1} \leftarrow \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (5)$$

where  $\hat{u}_k^{n+1}$  are obtained by Wiener filtering.

$$\lambda^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + \tau \left( \hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right) \quad (6)$$

The stopping criterion is formulated as follows:

$$\sum_k \frac{\|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \varepsilon \quad (7)$$

where,  $\tau$  is noise tolerance, and  $\varepsilon$  is convergence error.

### B. Pearson Correlation Coefficient

Correlation between two signals  $A$  and  $B$  of size  $N$  is the measure of their linear dependence. It is positively or negatively assessed if the correlation coefficient  $\rho$  is close to 1 or -1, respectively [31]. Pearson correlation coefficient is calculated as:

$$\rho(A, B) = \frac{1}{N-1} \sum_{i=1}^N \left( \frac{A_i - \mu_A}{\sigma_A} \right) \left( \frac{B_i - \mu_B}{\sigma_B} \right) \quad (8)$$

where  $\mu_A$ ,  $\sigma_A$ ,  $\mu_B$ , and  $\sigma_B$  are mean and standard deviations of  $A$  and  $B$ , respectively.

### C. Dominant Mode Filtering

This technique starts by decomposing signal  $x_n(t)$  into  $n$  modes, where at least one of them is closer to the original signal and called the dominant mode  $Mode_d$ . In this case, signal  $x_n(t)$  can be expressed as [31]:

$$x_n(t) = \sum_{i=1}^{d-1} Mode_i(t) + Mode_d(t) + \sum_{i=d+1}^n Mode_i(t) + res(t) \quad (9)$$

where  $d$  in  $Mode_d$  refers to dominant between the  $n$  modes. Location of the dominant mode is of major interest, especially for bearing fault early detection, as the sensed signal is usually dominated by other faults or even shaft rotating frequencies as previously shown in [32], [33]. Hence the need to eliminate this mode to keep a filtered signal  $x_c(t)$  containing only information related to the bearing fault. Dominant mode determination for elimination purpose is therefore an important step to increase bearing fault detection accuracy.

The filtered signal is the result of:

$$x_c(t) = x_n(t) - Mode_d \quad (10)$$

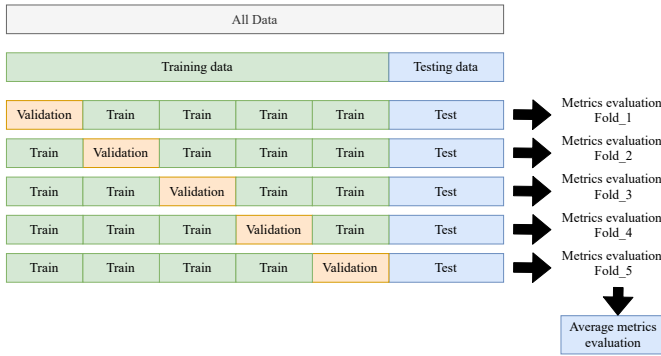


Fig. 2. Holdout test, 5-folds cross validation procedure [38].

#### D. Convolution Neural Network

CNN design is mainly based on the convolution of inputs with filters of different sizes to generate more discriminating output features, which will be inputs of the next layer. Pooling layers (Max, Average, or L2-norm) allow information compression and complexity reduction [9], in addition to overfitting control ensuring a better learning [34]. Convolution between input features  $u$  and Kernel filters  $k$  is provided by:

$$f = \varphi(u * k + b) \quad (11)$$

where  $f$  represents the obtained new features,  $*$  denotes the convolution operator,  $b$  is the bias, and  $\varphi$  is the activation function [35].

Feature extraction convolution and pooling layers are followed by feature learning layers (fully connected layers), which are traditional neural networks with an input, hidden, and classification layers [34].

#### E. Evaluation and Classification

Learning reliability is ensured by random sub-sampling technique [36]. The dataset is split randomly into subsets of training and testing data. Training data are then divided into training and validation for k-folds cross-validation technique, as shown in Fig. 2, to ensure a better reliability of prediction results. Evaluation is made according to universal metrics, such as accuracy [37] and confusion matrix.

### III. EXPERIMENTAL EVALUATION AND VALIDATION

#### A. Experimental Dataset

The proposed VMD/1D-CNN fault detection and diagnosis methodology is evaluated using the well-known Case Western Reserve University (CWRU) database [39] that has been extensively exploited in the literature for validation purposes [4], [10], [22].

The CWRU experimental setup, shown in Fig. 3, consists in a 1.49kW (2HP) three-phase Reliance electric motor driving a shaft on which a torque transducer and encoder are mounted. The used equipment includes an induction motor, a loading motor, and an axle attached to it.

The investigated bearing (SKF deep-groove ball bearings:

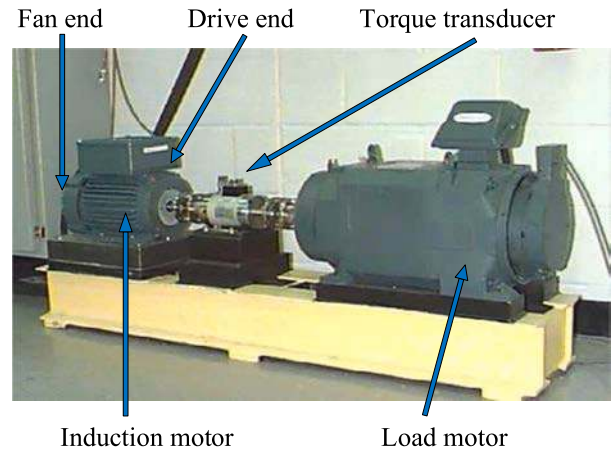


Fig. 3. CWRU test bench. Fig. 5. Bearing fault simulation test-bed.

TABLE I  
BEARING FAULT TYPES.

| Fault type   | Class label |
|--|-------------|
| Ball fault with a size of 0.007inch                | C1          |
| Ball fault with a size of 0.014inch                | C2          |
| Ball fault with a size of 0.021inch                | C3          |
| Ball fault with a size of 0.028inch                | C4          |
| Inner race fault with a size of 0.007inch          | C5          |
| Inner race fault with a size of 0.014inch          | C6          |
| Inner race fault with a size of 0.021inch          | C7          |
| Inner race fault with a size of 0.028inch          | C8          |
| Outer race fault centered with size of 0.007inch   | C9          |
| Outer race fault orthogonal with size of 0.007inch | C10         |
| Outer race fault opposite with size of 0.007inch   | C11         |
| Outer race fault centered with size of 0.014inch   | C12         |

6205-2RS JEM and 6203-2RSJEM) real faults concern inner race, outer race, and rolling elements with different severity and sizes (0.007, 0.014, 0.021, and 0.028inch). The CWRU dataset is enriched with 4 operating modes (0, 1, 2, and 3HP) at a speed of 1730rpm, with a sampling frequency of 12kHz. The proposed methodology is tested on drive-end accelerometer data for 12 different faults with different severities: Ball (0.007, 0.014, 0.021, and 0.028inch), inner race (0.007, 0.014, 0.021, and 0.028inch), outer race (centered, orthogonal, and opposite) for 0.007inch, and outer race centered for 0.014inch. These are summarized in Table I.

As illustrated in the methodology flowchart (Fig. 1), the above-presented data are pre-processed, i.e. segmented. Pre-processing consists in this case in data segmentation with overlap and window length of 8192 samples for each sub-signal. This is a sufficient length to maintain bearing faults features while increasing the number of sequences, resulting in a total of 28 sequences per fault.

#### B. Signal Processing

In this step VMD is used as a notch filter for dominant mode removing [25], [31]. In this study, decomposition consists in a maximum of 8 modes (Fig. 4) with center frequencies and bandwidth illustrated in Fig. 5.

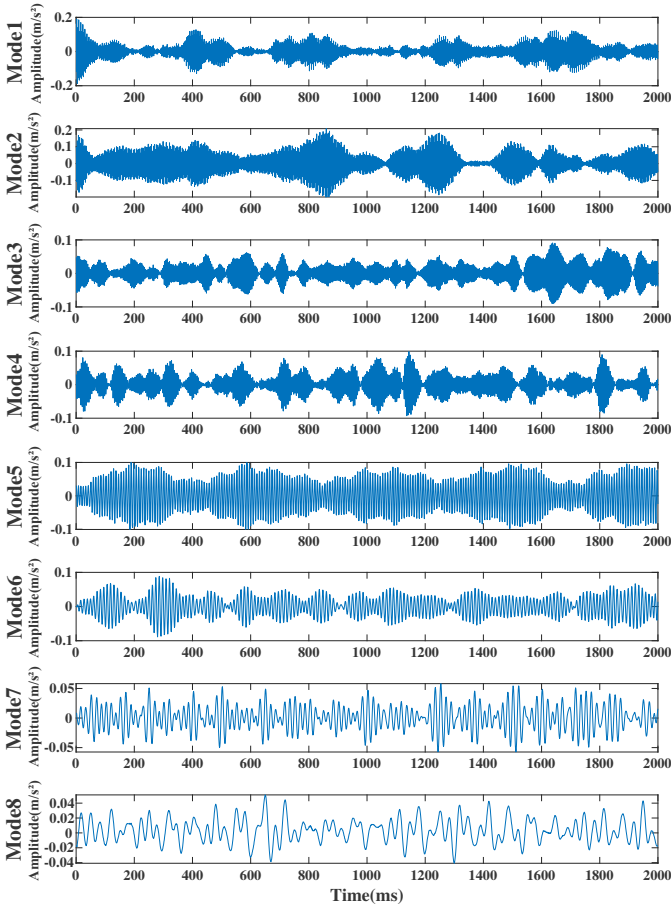


Fig. 4. Ball fault case VMD decomposition.

In this context, dominant mode identification is carried out using Pearson correlation between each decomposed mode and the original signal  $x_n(t)$ . The dominant mode (best correlation with  $\rho = 1$ ) is then subtracted from original signal leading to the filtered signal  $x_c(t)$  that will thereafter be used for detection purposes.

### C. Bearing Fault Detection and Diagnosis

Fault detection is in fact a binary classification between healthy and faulty states using 1D-CNN. In this context, VMD-based filtered signals  $x_c(t)$  first undergo treatment consisting in splitting signal into training and testing data with 70% and 30% for each class, respectively. Data-augmentation is then carried out with Additive White Gaussian Noise (AWGN) [40], with different Signal to Noise Ratio (SNR) levels (10, 15, 20, 25, 30, 35, and 40 dB), as shown in Fig. 6. The data preparation step is important as it allows improving learning quality by increasing the number of data, while improving fault detection robustness in a noisy environment [41].

For diagnosis purposes, the same 1D-CNN network is used with a difference in the last classification layer as illustrated by Fig. 7. For fault detection, the last layer allows a binary classification (healthy or faulty state). When a fault is detected, the fault diagnosis process is switch-on (event-triggered process).

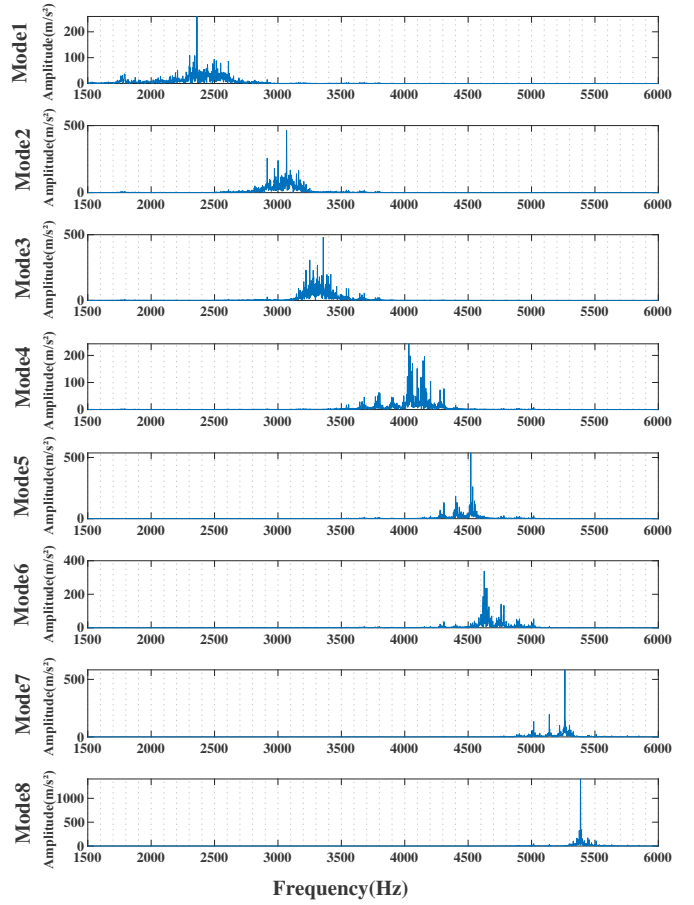


Fig. 5. Central frequencies and bandwidth of VMD decomposition.

Figs 8 and 9 illustrate the adopted CNN architecture for multi-scale features extraction ensuring fault detection and classification. This architecture is manually designed and tuned and consists in: (1) two 1D-convolution layers with 64 filters and kernel size of 10, (2) a dropout layer to control overfitting with a rate of 0.5 [42], (3) a max-pooling layer with pool-size of 2 for down-sampling and compressing useful information [22], and (4) a flattening layer to arrange features into vector. Features learning for classification uses a fully connected network of 300, 200, and 100 nodes for each layer, respectively, with a ReLU activation function and ends with a classification layer Softmax using a trial and error search technique. The training, tuned by a grid search mechanism [40], is done under the CPU with an early stopping option and batch size equal to 10 samples using Adam optimizer [43].

## IV. ACHIEVED RESULTS ANALYSIS AND DISCUSSION

Bearing faults detection has been achieved with 100% accuracy for the four operating modes (0, 1, 2, and 3HP). The proposed fault detection method was successful and efficient in discriminating healthy and faulty states. This is mainly due to the VMD use, as a notch filter removing the dominant mode that is common to healthy and faulty states, and the 1D-CNN use, allowing good discriminant features extraction. In terms of diagnosis, 5-folds cross validation is used. The VMD choice relevance, in addition to the above-mentioned

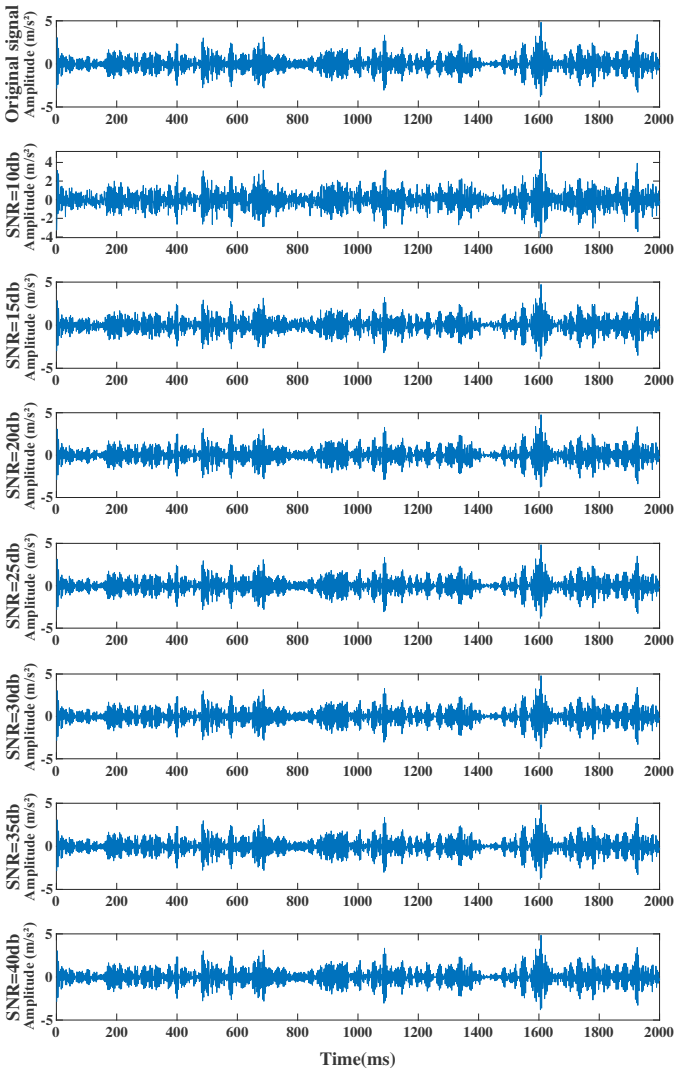


Fig. 6. AWGN with different SNR levels.

TABLE II  
PROCESSING TECHNIQUES ACCURACY COMPARISON.

|        | 0 HP          | 1 HP          | 2 HP          | 3 HP          |
|--------|---------------|---------------|---------------|---------------|
| EMD    | 67.04%        | 75.74%        | 78.70%        | 71.30%        |
| VMD    | 80.93%        | 75.56%        | 83.70%        | 79.44%        |
| VMD(-) | <b>91.30%</b> | <b>95.37%</b> | <b>93.15%</b> | <b>92.22%</b> |

validation, is evaluated versus the EMD, which is a similar adaptive time-frequency analysis technique but without a clearly mathematical basement [44]. Moreover, dominant mode filtering impact on detection accuracy is also evaluated (noted VMD(-)) versus a classical VMD (without dominant mode subtraction). The carried out comparison study clearly shows that the VMD(-) approach outperforms the two others in terms of accuracy, as illustrated by Table II, even under different working loads. Fault diagnosis performances are also highlighted by Figs. 10 to 13 in terms of confusion matrix.

Convolution layers (1D-CNN) learning relevance is also evaluated versus well-known machine learning approaches, namely multi-layer perceptron (MLP) and RNN, which are a simple networks without convolution layers, and LSTM, which

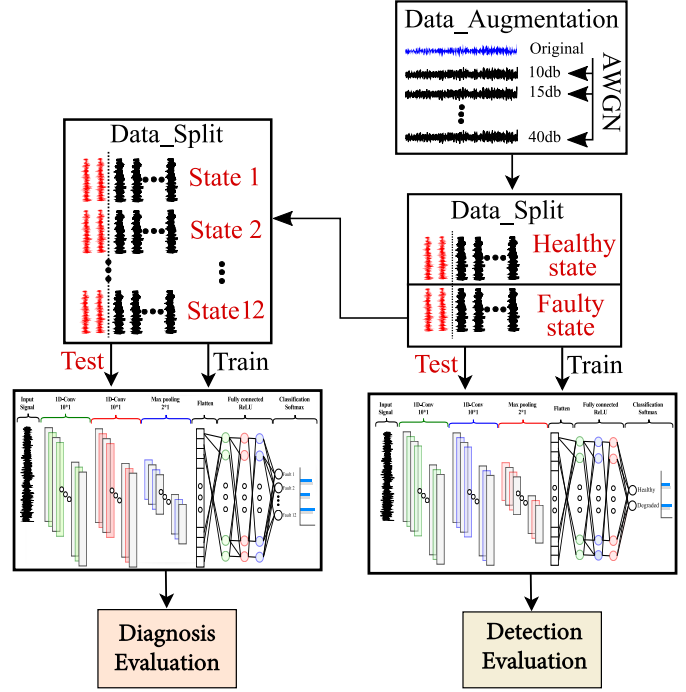


Fig. 7. Faults detection and diagnosis.

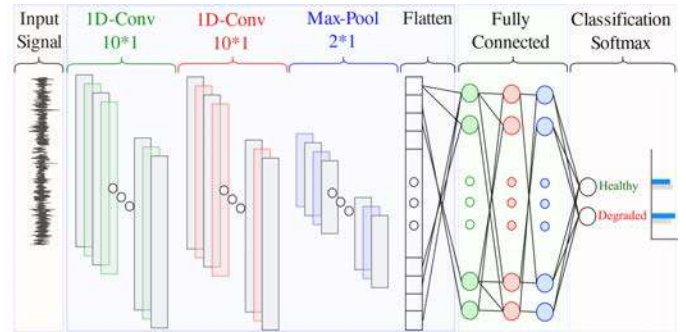


Fig. 8. Proposed 1D-CNN network for detection.

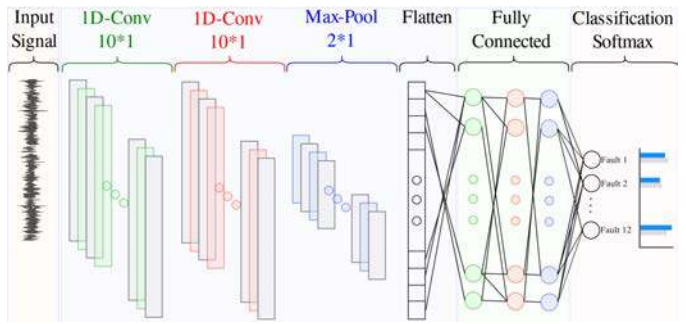


Fig. 9. Proposed 1D-CNN network for diagnosis.





severities. The detection and diagnosis performance have also been compared to the relevant state of the art corresponding techniques. In terms of fault detection, it has been shown that the VMD used as a notch filter is the solution of choice when compared to EMD and traditional VMD. In terms of features extraction for classification, convolution layers (1D-CNN) have been found worthwhile when compared with MLP, RNN, and LSTM.

Future investigations should consider exploring the proposed VMD-based 1D-CNN approach using LSTM for bearing failure prognosis. Indeed, LSTM is dedicated to time-series monitoring and clearly exhibit high learning ability of long- and short-terms dependencies, which can be useful for prognosis.

## REFERENCES

- [1] D. Wang, K.-L. Tsui, and Q. Miao, "Prognostics and health management: A review of vibration based bearing and gear health indicators," *IEEE Access*, vol. 6, pp. 665–676, 2018.
- [2] M. Benbouzid, Ed., *Signal Processing for Fault Detection and Diagnosis in Electric Machines and Systems*. Institution of Engineering and Technology, dec 2020.
- [3] C. Delpha, D. Diallo, J. Harmouche, M. Benbouzid, Y. Amirat, and E. Elbouchikhi, "Bearing fault diagnosis in rotating machines," *Electrical Systems 2: From Diagnosis to Prognosis*, pp. 123–151, 2020.
- [4] J. Gu, Y. Peng, H. Lu, S. Cao, and B. Cao, "Fault diagnosis of spindle device in hoist using variational mode decomposition and statistical features," *Shock and Vibration*, vol. 2020, 2020.
- [5] A. Mohammed and S. Djurovic, "Electric machine bearing health monitoring and ball fault detection by simultaneous thermo-mechanical fibre optic sensing," *IEEE Transactions on Energy Conversion*, pp. 1–1, 2020.
- [6] F. B. Abid, S. Zgarni, and A. Braham, "Distinct bearing faults detection in induction motor by a hybrid optimized SWPT and aiNet-DAG SVM," *IEEE Transactions on Energy Conversion*, vol. 33, no. 4, pp. 1692–1699, dec 2018.
- [7] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mechanical Systems and Signal Processing*, vol. 107, pp. 241–265, 2018.
- [8] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mechanical Systems and Signal Processing*, vol. 138, p. 106587, 2020.
- [9] J. Chen, W. Hu, D. Cao, M. Zhang, Q. Huang, Z. Chen, and F. Blaabjerg, "Novel data-driven approach based on capsule network for intelligent multi-fault detection in electric motors," *IEEE Transactions on Energy Conversion*, pp. 1–1, 2020.
- [10] C. E. Khamoudj, F. Benbouzid-Si Tayeb, K. Benatchba, M. Benbouzid, and A. Djaafri, "A learning variable neighborhood search approach for induction machines bearing failures detection and diagnosis," *Energies*, vol. 13, no. 11, p. 2953, 2020.
- [11] D. P. Mandic, N. u. Rehman, Z. Wu, and N. E. Huang, "Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 74–86, November 2013.
- [12] Y. Amirat, E. Elbouchikhi, Z. Zhou, M. Benbouzid, and G. Feld, "Variational mode decomposition-based notch filter for bearing fault detection," in *IECON 2019-45th Annual Conference of the IEEE Industrial Electronics Society*, vol. 1. IEEE, 2019, pp. 6028–6033.
- [13] J. B. Ali, L. Saidi, S. Harrath, E. Bechhoefer, and M. Benbouzid, "Online automatic diagnosis of wind turbine bearings progressive degradations under real experimental conditions based on unsupervised machine learning," *Applied Acoustics*, vol. 132, pp. 167–181, 2018.
- [14] S. Tang, S. Yuan, and Y. Zhu, "Deep learning-based intelligent fault diagnosis methods toward rotating machinery," *IEEE Access*, vol. 8, pp. 9335–9346, 2019.
- [15] G. Cirrincione, R. R. Kumar, A. Mohammadi, S. H. Kia, P. Barbiero, and J. Ferretti, "Shallow versus deep neural networks in gear fault diagnosis," *IEEE Transactions on Energy Conversion*, vol. 35, no. 3, pp. 1338–1347, sep 2020.
- [16] T. Xia, Y. Song, Y. Zheng, E. Pan, and L. Xi, "An ensemble framework based on convolutional bi-directional LSTM with multiple time windows for remaining useful life estimation," *Computers in Industry*, vol. 115, p. 103182, feb 2020.
- [17] X. Chen, Y. Yang, Z. Cui, and J. Shen, "Vibration fault diagnosis of wind turbines based on variational mode decomposition and energy entropy," *Energy*, vol. 174, pp. 1100–1109, 2019.
- [18] Y. Li, G. Cheng, C. Liu, and X. Chen, "Study on planetary gear fault diagnosis based on variational mode decomposition and deep neural networks," *Measurement*, vol. 130, pp. 94–104, 2018.
- [19] J. Gai, J. Shen, Y. Hu, and H. Wang, "An integrated method based on hybrid grey wolf optimizer improved variational mode decomposition and deep neural network for fault diagnosis of rolling bearing," *Measurement*, vol. 162, p. 107901, 2020.
- [20] V. Sharma and A. Parey, "Extraction of weak fault transients using variational mode decomposition for fault diagnosis of gearbox under varying speed," *Engineering Failure Analysis*, vol. 107, p. 104204, 2020.
- [21] Z. Wang, Y. Dong, W. Liu, and Z. Ma, "A novel fault diagnosis approach for chillers based on 1-d convolutional neural network and gated recurrent unit," *Sensors*, vol. 20, no. 9, p. 2458, 2020.
- [22] S. Hao, F.-X. Ge, Y. Li, and J. Jiang, "Multisensor bearing fault diagnosis based on one-dimensional convolutional long short-term memory networks," *Measurement*, p. 107802, 2020.
- [23] X. Liu, Q. Zhou, J. Zhao, H. Shen, and X. Xiong, "Fault diagnosis of rotating machinery under noisy environment conditions based on a 1-d convolutional autoencoder and 1-d convolutional neural network," *Sensors*, vol. 19, no. 4, p. 972, 2019.
- [24] G. Jiang, H. He, J. Yan, and P. Xie, "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3196–3207, 2018.
- [25] A. Dibaj, M. M. Etefagh, R. Hassannejad, and M. B. Eghghaghi, "A hybrid fine-tuned VMD and CNN scheme for untrained compound fault diagnosis of rotating machinery with unequal-severity faults," *Expert Systems with Applications*, p. 114094, oct 2020.
- [26] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, feb 2014.
- [27] X. Chen, Y. Yang, Z. Cui, and J. Shen, "Vibration fault diagnosis of wind turbines based on variational mode decomposition and energy entropy," *Energy*, vol. 174, pp. 1100 – 1109, 2019.
- [28] Z. Wang, G. He, W. Du, J. Zhou, X. Han, J. Wang, H. He, X. Guo, J. Wang, and Y. Kou, "Crude oil risk forecasting: New evidence from multiscale analysis approach," *IEEE Access*, vol. 7, pp. 44 871–44 882, 2019.
- [29] F. Jiang, Z. Zhu, and W. Li, "An improved VMD with empirical mode decomposition and its application in incipient fault detection of rolling bearing," *IEEE Access*, vol. 6, pp. 44 483–44 493, 2018.
- [30] R. Gu, J. Chen, R. Hong, H. Wang, and W. Wu, "Incipient fault diagnosis of rolling bearings based on adaptive variational mode decomposition and teager energy operator," *Measurement*, vol. 149, p. 106941, 2020.
- [31] Y. Amirat, M. Benbouzid, T. Wang, K. Bacha, and G. Feld, "EEMD-based notch filter for induction machine bearing faults detection," *Applied Acoustics*, vol. 133, pp. 202–209, apr 2018.
- [32] X. Zhang, J. Kang, L. Xiao, J. Zhao, and H. Teng, "A new improved kurtogram and its application to bearing fault diagnosis," *Shock and Vibration*, vol. 2015, 2015.
- [33] W. Wang and H. Lee, "An energy kurtosis demodulation technique for signal denoising and bearing fault detection," *Measurement Science and Technology*, vol. 24, no. 2, p. 025601, 2013.
- [34] C. Wu, P. Jiang, C. Ding, F. Feng, and T. Chen, "Intelligent fault diagnosis of rotating machinery based on one-dimensional convolutional neural network," *Computers in Industry*, vol. 108, pp. 53–61, 2019.
- [35] X. Li, W. Zhang, and Q. Ding, "Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction," *Reliability Engineering & System Safety*, vol. 182, pp. 208–218, 2019.
- [36] M. Xia, T. Li, L. Xu, L. Liu, and C. W. De Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 1, pp. 101–110, 2017.
- [37] M. T. Pham, J.-M. Kim, and C. H. Kim, "Deep learning-based bearing fault diagnosis method for embedded systems," *Sensors*, vol. 20, no. 23, p. 6886, 2020.
- [38] Z. Xiong, Y. Cui, Z. Liu, Y. Zhao, M. Hu, and J. Hu, "Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation," *Computational Materials Science*, vol. 171, p. 109203, 2020.
- [39] K. Loparo, "Bearing data center seeded fault test data," *The Case Western Reserve University Bearing Data Center Website [EB/OL]*, 2005.

- [40] T. Berghout, L.-H. Mouss, T. Bentrchia, E. Elbouchikhi, and M. Benbouzid, "A deep supervised learning approach for condition-based maintenance of naval propulsion systems," *Ocean Engineering*, vol. 221, p. 108525, 2021.
- [41] Z. Xu, C. Li, and Y. Yang, "Fault diagnosis of rolling bearing of wind turbines based on the variational mode decomposition and deep convolutional neural networks," *Applied Soft Computing*, vol. 95, p. 106515, 2020.
- [42] M. Roder, G. H. de Rosa, V. H. C. de Albuquerque, A. L. Rossi, and J. P. Papa, "Energy-based dropout in restricted boltzmann machines: Why not go random," *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2020.
- [43] J. Chi, X. Li, H. Wang, D. Gao, and P. Gerstoft, "Sound source ranging using a feed-forward neural network trained with fitting-based early stopping," *The Journal of the Acoustical Society of America*, vol. 146, no. 3, pp. EL258–EL264, 2019.
- [44] Y. Amirat, V. Choqueuse, and M. Benbouzid, "EEMD-based wind turbine bearing failure detection using the generator stator current homopolar component," *Mechanical Systems and Signal Processing*, vol. 41, no. 1-2, pp. 667–678, dec 2013.
- [45] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mechanical Systems and Signal Processing*, vol. 100, pp. 439–453, 2018.