Firas Ben Abid[1], **Marwen Sallem**[1], **Ahmed Braham**[1, 2*]

J. Electrical Systems 18-4 (2022): 446-460

Regular paper

An End- To-End Bearing Fault Diagnosis and Severity Assessment with Interpretable Deep Learning

The induction motor (IM) consumes 85% of the electrical energy in industry applications. The crucial place occupied by the IM in the industry requires the installation of an efficient integrated machinery health management (MHM) process in order to improve the availability of this equipment and minimize the cost of maintenance. Characterized by a high potential of unsupervised feature learning, the Deep Learning (DL)-based MHM of IM has recently been introduced. However, it is still hard to interpret physically the features learned by most of the already implemented deep architectures. In this paper, an architecture called Deep-SincNet is implemented for bearing faults (BF) diagnosis task. The proposed scheme automatically learns the fault features from the raw vibration signals and accordingly finalizes the fault diagnosis process. As well as the features related to BF characteristic frequencies, the proposed architecture automatically extracts further interpretable features. The performance of the proposed approach is tested for two different databases. For several types of bearing faults, different levels of fault severity and variable loads, the Deep-SincNet provides high accuracy, a significant gain in implementation cost, faster convergence, and interpretable results.

Keywords: Condition Monitoring; Convolutional Neural Network; Fault Diagnosis; Deep Learning; Induction Motor.

1. Introduction

In all industrial applications, induction motors (IM) represent a key part of the process. Shortly, the global market of IM is estimated to reach 22.7 billion dollars, while it was valued at 12.7 billion dollars in 2017 [1]. The robustness, the low cost, and the environmental protection make IM the most appropriate choice for manufacturing industries. However, IMs are submitted to hostile conditions, so they are commonly affected by faults. The maintenance of industrial machinery has always represented a critical need. As industrial systems' complexity increased, the maintenance policies evolved. These latter started with reliability analysis in the 1960s, then they evolved to Condition-Based Maintenance in the 1990s and by the 2000s, they were extended to System Health Management and Machinery Health Management (MHM) [2]. The MHM necessarily includes an anomaly detection mechanism for fault detection and a diagnosis process for the identification and severity assessment of the fault [3].

Several techniques were applied for the fault detection and diagnosis of IM faults. Since the majority of IM defects occur in bearings (40% to 90% from high power to low power IM) [4], most of the works focused on bearing fault (BF) detection and diagnosis. The BF is

* Corresponding author: A. Braham, University of Carthage, Tunis 1054, Tunisia,
Matériaux, Mesures et Applications (MMA) Laboratory, INSAT, Tunis 1080, Tunisia,
E-mail: ahmed.braham@insat.rnu.tn
[1] Matériaux, Mesures et Applications (MMA) Laboratory, INSAT, Tunis 1080, Tunisia
[2] University of Carthage, Tunis 1054, Tunisia

usually diagnosed by vibration analysis due to its direct correlation with the vibration of the motor. The vibration signals are commonly analyzed by spectral and spectral-temporal techniques. The envelope analysis, the Fast Fourier Transform, the Empirical Mode Decomposition, the Short Time Fourier Transform, the Variational Mode Decomposition, the Wavelet Transforms, and others performed efficient systems of BF detection and diagnosis [5]-[16].

Along with the current orientation toward industry 4.0, plant managers tend to reduce human intervention and guarantee that industrial applications are accomplished with more autonomy. Such trends led to the appearance of new maintenance policies such as self-maintenance, e-maintenance, and engineering immune systems [17]. The combination of Artificial Intelligence (AI)-based techniques, with the aforementioned signal processing approaches has guaranteed an automatic classification of the BF. Several classifiers were applied, such as Artificial Neural Networks [18], k-Nearest Neighbors [15], Random Forests [11], Support Vector Machines [19], Extreme Learning Machines [12], [13] and even Multiple Classifiers Fusion [14]. The use of the handcrafted features as inputs of the classifiers efficiently performed interpretable results and confident decision-making [5]-[16].

However, human intervention in feature design is still highly required even with introducing conventional AI-based classifiers. A lot of expertise in both fault diagnosis and signal processing is necessary [3], to first pre-process the signals, then design, optimize and extract the fault descriptors and finally eliminate the redundant data by feature selection algorithms, such as Principal Component Analysis and Neighbourhood Rough Set [5], [8], [13], [14]. Besides, the manual feature design is commonly based on the tracking of the fault characteristic frequencies. In practice, the appearance of the frequency components of some theoretical characteristic frequencies of the BF is not guaranteed [20].

In fact, at a very early stage, the BF does not excite any characteristic frequency, but it rather generates mid and high-range frequency components above 500 Hz [20]. Therefore, by limiting oneself to the surveillance of the characteristic frequencies, the detection and the identification of faults can be compromised. Moreover, a large quantity of the signatures generated by the faults is ignored [21].

Since machine learning-based architectures have become deeper, Deep Learning (DL) has attracted the attention of researchers in several domains including fault diagnosis. Several researchers have used handcrafted feature vectors as inputs of the DL algorithms, similarly to the conventional classifiers [22]-[24]. Different DL architectures have been combined with temporal, spectral, or spectro-temporal feature extraction techniques. Nevertheless, these approaches are not appropriate with the philosophy of DL, which opposes manual feature extraction. Actually, the DL provides a high capacity for automatic feature extraction, also known as unsupervised feature learning. It is designed for the automatic generation of unknown descriptors in the signals and to provide an end-to-end system [3].

To get back into the philosophy of the DL, recent works have completely eliminated the step of manual feature extraction [25]-[32]. Since the DL was originally applied in image processing, several researchers converted the vibration signals into 2-D vibration images [25]-[27]. In order to allow an on-line BF diagnosis, which tends to deal directly with raw data, further works have applied the 1-D DL. Several BF diagnosis frameworks have been

developed based on auto-encoders [28], [29], Deep Belief Networks (DBN) [30], Convolutional Neural Networks (CNN) [31], and Long Short-Term Memory networks [32].

However, most of the already implemented deep architectures are still considered a black box, which suffers from a lack of physical interpretability [3]. In this paper, the authors represent a new alternative for an end-to-end FD approach, which automatically and directly learns features of multiple BF from the raw vibration signal and jointly trains the system to the fault classification. In the first dataset, nine bearing conditions that include different bearing faults are diagnosed. A novel deep architecture called Deep-SincNet is implemented. The developed scheme includes a customized filter bank, which is automatically and precisely fitted with the application. Unlike the aforementioned deep architectures, the proposed approach merges the advantages of the conventional feature extraction techniques in terms of clear interpretability with the advantages of the DL in terms of unsupervised feature learning. The proposed methodology performs more physical interpretability, a faster convergence, and a lesser consumption in terms of implementation cost. In order to provide a clear comparison with previous methodologies, the effectiveness of the Deep-SincNet is then tested on the Case Western Reserve University (CWRU) database. The Deep-SincNet proves once more its efficiency and provides a high fault diagnosis accuracy.

The main contributions of this work consist of the following points:

- The application of a new DL architecture directly on the raw vibration signals to diagnose multiple BF and assess their severity;
- The implementation of an interpretable DL architecture, which extracts clear features, converges faster and requires lesser computational cost;
- The recognition of new fault patterns by the interpretation of the feature extraction process of the Deep-SincNet;
- The achievement of high precision for the detection of the anomalies, the recognition of the fault type, and the identification of the severity level of the same BF;
- The achievement of higher accuracy compared to previous works on a benchmark database.

The organization of the rest of the paper is as follows: Section 2 represents the Deep-SincNet architecture. Section 3 defines the experiments conducted to validate the proposed approach, while the results are represented and discussed in section 4. Finally, section 5 concludes the paper with perspectives and highlights.

2. The Deep-SincNet Architecture

Generally, the extraction of descriptors is demanding in terms of expertise in both fault diagnosis and signal processing, although it may lead to unstable performance levels [3]. The DL offers at this stage a high aptitude to overcome the requirements of conventional AI methods that must be fed by handcrafted fault descriptors. CNN has been frequently used in several domains and it has been successfully applied in fault diagnosis [31]. CNN is a deep neural network composed of several filtering steps and a classification step. Its learning process comprises an automatic feature extractor. This extraction system is based on layers of convolutions, activation layers, and pooling layers. The processing of data by the convolution process is defined by the equation (1) [31]:

$$y^{l(i,j)} = K_i^l * x^{l(r^j)} = \sum_{k=0}^{M} K_i^l(k).x^{l(j+k)} \quad (1)$$

where K_i^l is the weight of the kernel of i the filter in a layer l, $x^{l(r^j)}$ denotes the j^{th} local region in the layer *l*, * is the convolution product, *M* represents the kernel width, and $K_i^l(k)$is the k^{th} kernel weights. In order to speed up the learning process and reduce the inherent shift covariance, a batch normalization layer is introduced. This latter is followed by a Rectified Linear Unit (ReLU) that offers a better Gradient propagation while adjusting the parameters by the method of back-propagation. The outputs of the convolution layers are reduced in dimensions by using pooling layers. Finally, a fully connected layer classifies the feature vectors selected from the previous steps and generates the outputs. However, like other DL architectures, the CNN is still considered a "black box" whose physical interpretation remains difficult [3]. Although the filters learned by the first convolution layer have a very important role in the CNN architecture, they remain difficult to interpret and their spectral content is highly noisy. In order to overcome this disadvantage, SincNet was recently developed by Ravanelli and Bengio [33] to make CNN interpretable. Unlike conventional CNN, the SincNet architecture strengthens the first convolutional layer by creating a set of temporal convolutions between the input raw signal and interpretable digital filters. The implementation of SincNet is based on cardinal-sinus functions (Sinc) for the design of bandpass filters. Equation (2) defines each convolution:

$$y[n] = x[n] * h[n] = \sum_{l=0}^{L-1} x[l] \cdot h[n-l] \quad (2)$$

Whereas the conventional CNN learns all the parameters of the applied filters, the Sinc convolutions are generated by a predefined function g, which allows for reducing the number of parameters. This principle is expressed by equation (3):

$$y[n] = x[n] * g[n, \theta] \quad (3)$$

A physically interpretable filter bank is therefore designed and adapted exactly to each application. The use of rectangular filters is a judicious choice for developing an adjustable filter bank. From the input signal, the Sinc filters only learn the cutoff frequencies specific to each filter. The spectral representation of the filters learned by SincNet is defined by equation (4):

$$G[f, f_L, f_H] = rect\left(\frac{f}{2f_H}\right) - rect\left(\frac{f}{2f_L}\right) \quad (4)$$

where f_L and f_H are respectively the low and high cutoff frequencies learned by the Sinc filters and *rect* represents the rectangular function. The temporal representation of the Sinc filters becomes as follows:

$$g[n, f_L, f_H] = 2f_H \; sinc(2\pi f_H n) - 2f_L \, sinc(2\pi f_L n) \quad (5)$$

In order to idealize the spectral shape of the filters using a finite number of parameters, the filters are windowed by $w[n]$. This principle is expressed by equation (6):

$$g_w[n, f_L, f_H] = g[n, f_L, f_H].w[n] \quad (6)$$

Similar results are obtained regardless of the type of the applied window, namely, Hamming, Blackman, Hanning or Kaiser [33]. Thus, the Sinc convolutional layer designs a compact, robust and interpretable filter bank. The rest of the process can be provided by the conventional CNN architecture, which includes layers of pooling and a softmax classification.

3. Description of the experimental bench and the sincNet model

3.1. Experimental bench

As exhibited in Fig. 1, the experiments are realized by a 3 kW IM with 4 poles, equipped with 6206 ball bearings. The IM is coupled to a DC generator in order to allow the load level variation [4], [34]. A low-cost IEPE accelerometer PCB 602D01 and an NI 9234 data acquisition board acquire the vibration signals. The tests are about healthy IM (HLT) and eight shaft-end bearing faults. The bearing faults include three outer race faults and three inner race faults with different severities, a cage fault and a ball fault. In order to test the sensitivity of the technique to different levels of severities holes of 0.5 mm, 1 mm and 3 mm are drilled using the electric discharge machining in the outer race and the inner race of several bearings, as shown in Fig. 2. The cage of another bearing is partially broken to create the cage fault (CF). For the ball fault, a pitting is artificially created by the wire-cut electric discharge machining, which is dedicated to machine very delicate and intricate shapes. Nine motor conditions are therefore tested. Table 1 regroups all the diagnosed bearing conditions. For the nine motor conditions, the vibration signals are captured at 0%, 25%, 50%, 75% and 100% of the load. At each load level, 30 acquisitions for each motor condition are collected with a sampling frequency of 25.6 kHz and a number of samples of 128 kS.

Unlike the conventional handcrafted feature extraction frameworks that usually exploit the entire input signal, the proposed approach only deals with several short-time segments.These latter are built by the sliding window technique, which technically splits the signal $x[n]$ into several segments with equal sizes S_1, S_2,….. S_M. In the present application, each window S_i has a length of m = 12.8 kS, which corresponds to 0.5 seconds of the vibration signal. Therefore, the number of inputs of the deep learning model is 10 times increased. As proved by the experiments, the 0.5-second segment is sufficient to extract the fault-related patterns.

Table 1: Motor conditions under test

Label	**HLT**	**BLF**	**CF**	**IRF**			**ORF**		
				IRF1	IRF2	IRF3	ORF1	ORF2	ORF3
Fault location	-	Ball	Cage	Inner race	Inner race	Inner race	Outer race	Outer race	Outer race
Fault Size (mm)	0			0.5	1	3	0.5	1	3

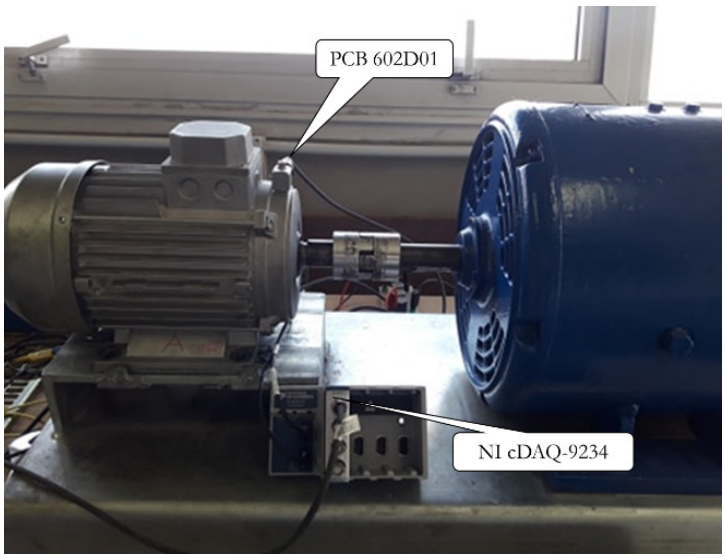
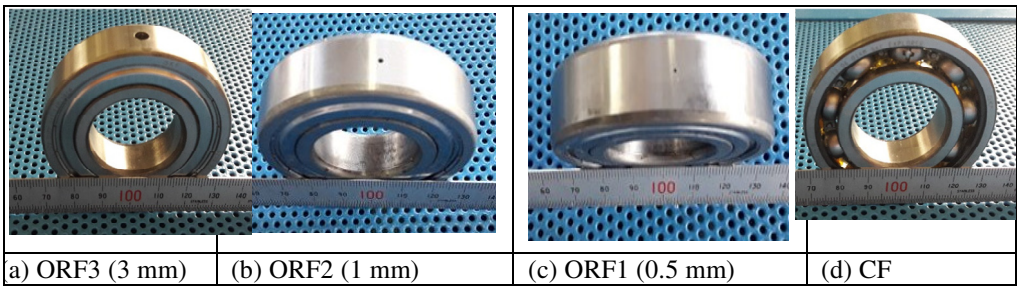
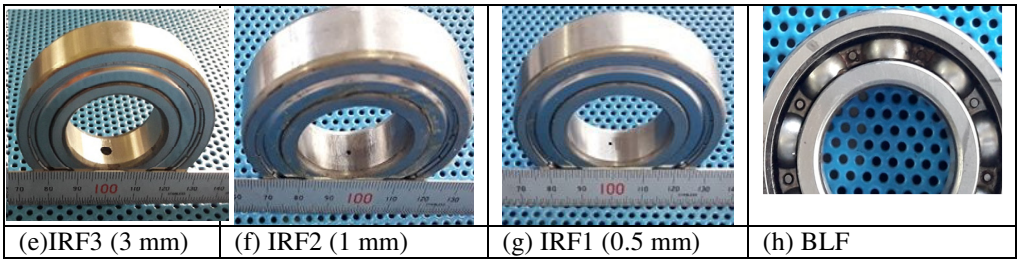


Fig. 1. Experimental bench for multi-fault diagnosis



a) ORF3 (3 mm)	(b) ORF2 (1 mm)	(c) ORF1 (0.5 mm)	(d) CF



(e)IRF3 (3 mm)	(f) IRF2 (1 mm)	(g) IRF1 (0.5 mm)	(h) BLF

Fig. 2. Bearing Faults under diagnosis

3.2. Design of the Deep-SincNet model

As shown in Table 2, the selected model is composed of a *Sinc*-convolutional layer, a standard convolutional layer, and a fully connected layer with softmax. A batch normalization layer and a maxpooling layer follow each convolutional layer. The generalization ability is still a major weakness of neural networks. This handicap leads to over-fitting problems. To overcome this disadvantage, a global average pooling is used just before the fully connected layer by reducing the total number of parameters of the model. To ensure the global average pooling, the Adam optimization algorithm minimizes the

Categorical Cross Entropy, where 0.9 and 0.999 respectively define $\beta 1$ and $\beta 2$ and the learning rate is set at 0.001. A Leaky ReLU is used as a nonlinear activation function for the first two convolutional layers.

The selected model offers the best compromise between accuracy, speed of classification, and cost of implementation compared to other architectures. The training task is performed by 3/4 of the data, while the others are used for the validation. The tests are executed via a 2.4 GHz i7 processor with 8 cores and equipped with an 8 GB memory.

Table 2: Model parameters

Layer	Hyper parameters	Output Shape	Number of parameters
Input layer	-	(12800, 1)	0
Sinc-convolutional layer	F = 50, L = 101	(12700, 50)	100
Batch Normalization	-	(12700, 50)	100
1-D MaxPooling	pooling size = 3	(4233, 50)	0
Leaky Relu	α = 0.2	(4233, 50)	0
1-D Convolutional Layer	F = 10 L = 20	(4214, 10)	10010
Batch Normalization	-	(4214, 10)	40
1-D Max Pooling	pooling size = 3	(1404, 10)	0
Leaky ReLu	α = 0.2	(1404, 10)	0
1-D Global Average Pooling	-	(10)	0
Fully connected layer with softmax	Number of units = 9 (number of classes)	(9)	99

4. Results and discussion

4.1. Unsupervised feature learning

In practice, most conventional diagnosis systems depend on the estimation of the characteristic frequencies of each type of defect [9]. The characteristic vibration frequencies related to the bearing faults are expressed by the following equations:

$$f_o = \frac{f_r}{2} N_b \left(1 - \frac{Ball\ \ Diameter}{Pitch\ \ Diameter} \cos \alpha\right) \quad (8)$$

$$f_i = \frac{f_r}{2} N_b \left(1 + \frac{Ball\ \ Diameter}{Pitch\ \ Diameter} \cos \alpha\right) \quad (9)$$

$$f_c = \frac{f_r}{2} \left(1 - \frac{Ball\ \ Diameter}{Pitch\ \ Diameter} \cos \alpha\right) \quad (10)$$

$$f_b = \frac{f_r}{2} \frac{Pitch\ Diameter}{Ball\ Diameter} \left(1 - \frac{Ball\ Diameter^2}{Pitch\ Diameter^2} \cos^2 \alpha\right) \quad (11)$$

where f_o, f_i, f_c and f_b denote the characteristic frequencies of the outer race, the inner race, the cage and the ball, f_r is the rotational mechanical frequency, N_b is the number of balls and α denotes the contact angle of the bearing.

However, some of these parameters are sometimes inestimable for already operating motors. In addition, the appearance of the frequency components of some theoretical characteristic frequencies of BF is not guaranteed [20], [21]. In fact, the degradation process of the bearing progresses through four stages of failure. During the first two stages, the very incipient bearing fault does not excite the aforementioned characteristic frequencies, but it

rather excites mid and high-range frequency components above 500 Hz that differ according to the motor power [20]. Therefore, only monitoring the characteristic frequencies can on the one hand compromise the detection and the identification of faults, and on the other hand ignores a large quantity of the signatures generated by the faults [21].

The proposed alternative offers a compact solution for using further fault descriptors that are still unknown and not exploited by the mathematical models already available. Fig. 3 illustrates the difference between the conventional frameworks and the present work. This new diagnosis system helps to escape from estimating the characteristic frequencies, resorting to preprocessing, or even manually extracting and selecting the fault descriptors.

The results from the proposed approach are illustrated in Fig. 4. The generated outputs are visualized using the *t*-distributed Stochastic Neighbor Embedding (t-SNE) algorithm. Developed by Geoffrey Hinton and Laurens van der Maaten, the latter is a technique of dimension reduction and data visualization. It presents a non-linear method that allows projecting a set of big data in a reduced space of two or three dimensions illustrated by a scatter plot.

The first-level scatter plot of Fig. 4 is related to the data extracted from the Sinc convolutional layer for the 9 bearing conditions. It can be observed that these samples are overlapped and difficult to distinguish. As the level of the layers becomes deeper, the clusters associated with each class become distinguishable. This phenomenon reveals the high performance of the deep architectures. A Sinc convolutional layer and a standard convolutional layer are sufficient to discriminate the 9 bearing conditions. In the last layer, the samples specific to each bearing condition are grouped together, even for different load levels. The SincNet shows high classification capability for the diagnosis of distinct BF, even at different levels of fault severity and at any load level. In fact, several descriptors characterizing each bearing condition are automatically extracted independently of the load level. To give further details of the diagnostic results, the confusion matrices resulting from the CNN and SincNet models are represented in Fig. 5. Table 3 summarizes the classification results by CNN and SincNet.

Table 3: Comparison of the performance betweenDeep-SincNet and CNN

Model	**Average testing accuracy**	**Area under curve ROC score**	**Number of epochs for data training**
CNN	98.21%	0.997	261
Deep-SincNet	100 %	1.000	39

The average precision rate achieved by SincNet is 100% and 98.21% by CNN. With CNN, it can be noted that misclassification occurs in the case of the incipient ORF. The classifier misclassifies 20% of the ORF1 samples as IRF2.

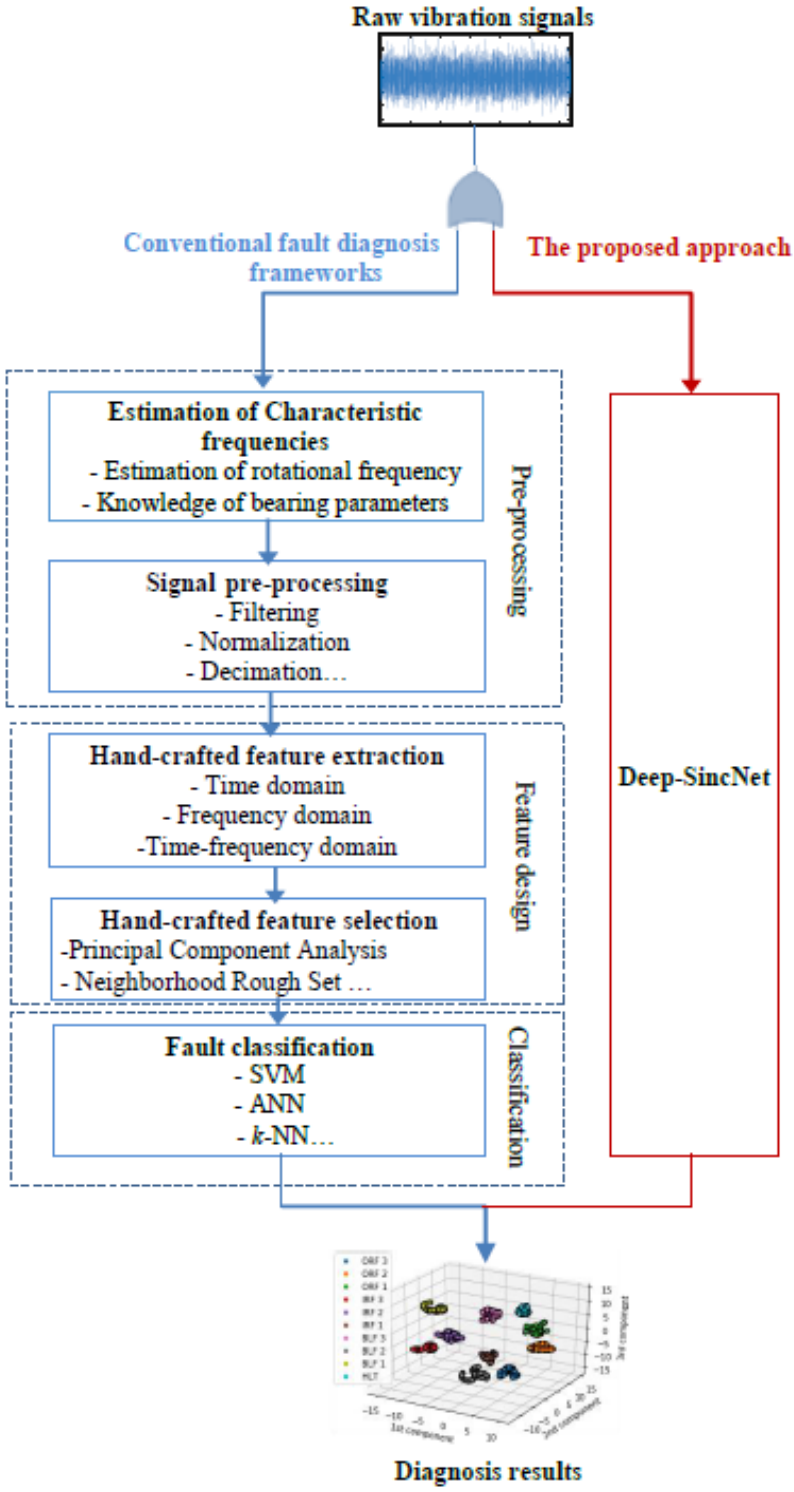


Fig. 3. Methodologies of the conventional fault diagnosis frameworks and the proposed approach

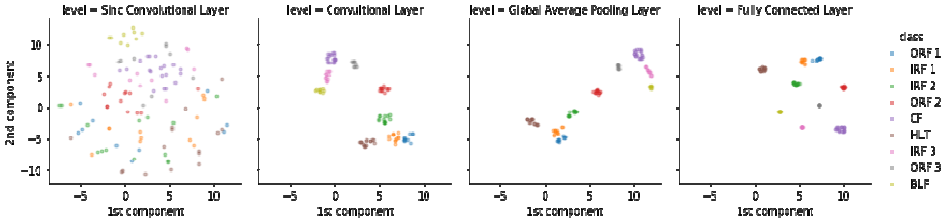
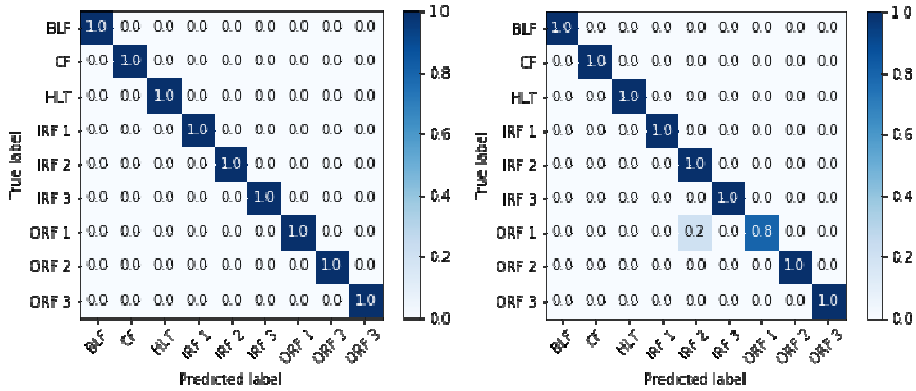


Fig. 4. t-SNE multi-class visualization



(a) Normalized confusion matrix by Deep-SincNet (b) Normalized confusion matrix by CNN

Fig. 5. Confusion matrices of multi-class classification

4.2. Fast Convergence

From Table 3, it can be also observed that the SincNet converges after a number of epochs practically seven times lesser than the CNN. Another advantage of the SincNet architecture lies in its fast convergence compared to CNN. SincNet focuses only on a few filtering parameters that have a major impact on performance. Like the conventional feature extraction methods, the SincNet uses well-known filters, while holding its flexibility to be adapted to each acquired signals. This prior knowledge of the parameters facilitates the learning of the characteristics of the filters, thus helping the SincNet to converge more quickly.

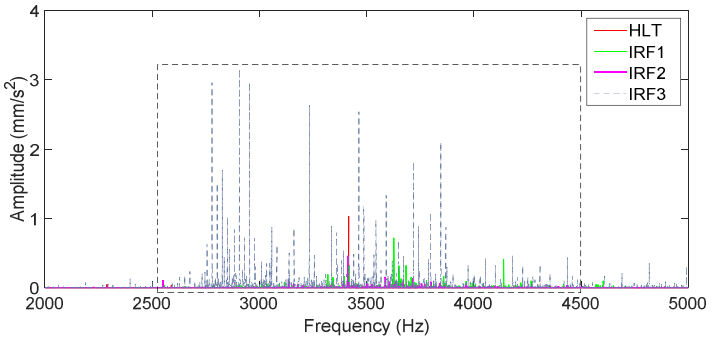
4.3. Interpretability

The proposed SincNet architecture makes an important contribution to the interpretation of the internal process of the DL. Unlike the kernels designed by CNN, those of the Sinc filters have a clear physical meaning. The design of the Sinc filter bank only requires a few parameters with a clear physical meaning, namely, the cutoff frequencies. The filters designed by SincNet are band-pass filters withrecognizable cut-off frequencies. On the other hand, the filters learned by CNN are deprived of any mathematical definition of their temporal or spectral responses and they suffer from noisy spectral content. The fact that the first layer of SincNet designs standard filters leads to the recognition of the frequency bands selected from the automatic feature-learning step.Table 4 exhibits examples of the frequency bands $[f_L - f_H]$ Hz extracted by SincNet.

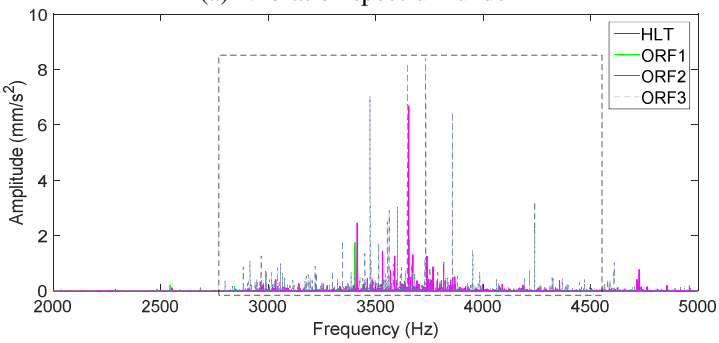
Table 4: Examples of the frequency bandwidths extracted by SincNet

f_L (Hz)	52.6	136.3	2427.6	2138.1	5292.8
f_H (Hz)	156.3	211.4	4193.4	4313.6	10488.0

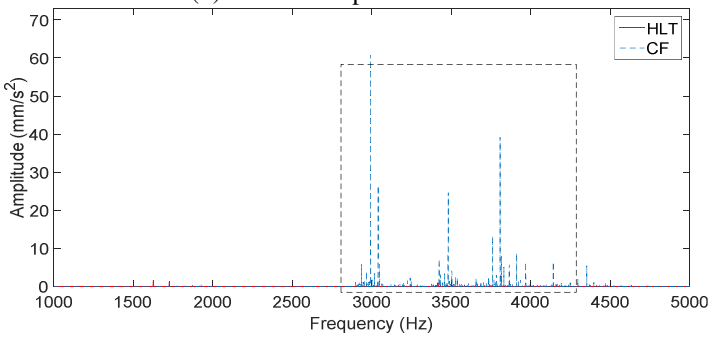
It can be emphasized that the Sinc filters extract the commonly used low-frequency fault descriptors. In addition, they generate additional descriptors in mid and high-range frequencies. On the one hand, the robustness of the network is subsequently reinforced against the effects of the dominant frequency components. On the other hand, it is clear that



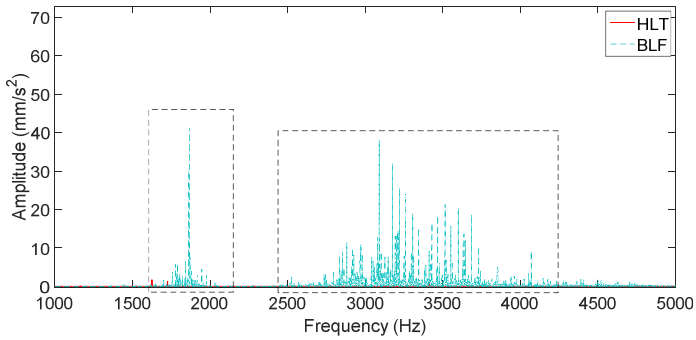
(a) Vibration spectrum under IRF



(b) Vibration spectrum under ORF



(c) Vibration spectrum under CF



(d) Vibration spectrum under BLF

Fig. 6. Vibration spectrums under the different bearing conditions at the mid-range frequencies extracted by SincNet (linear scale)

the SincNet architecture exploits the frequency descriptors excited in the first two stages of the bearing degradation. Although they represent very important fault descriptors especially at the incipient stages of the bearing failure development, these frequencies are generally ignored by the conventional approaches, while they are automatically extracted by the SincNet. The importance of these mid-range frequencies extracted by SincNet can be seen in the vibration spectrum illustrated by Fig. 6.

4.4. Reduction of implementation cost

In order to implement the intelligent diagnosis system on embedded systems, reducing implementation costs and storage space becomes the main concern. For a hardware implementation of a DL architecture, the number of parameters (NP) is an important handicap. As long as the architecture becomes deeper, the NP extremely increases. Thus, the size of the model rises and causes storage problems, especially when the model is implemented on low-cost devices. A low memory footprint characterizes the SincNet models since they use a reduced NP compared to CNN. Each layer is composed of F filters with a length of L. The CNN would need F × L parameters. However, a Sinc layer would only require F × 2 parameters (two cutoff frequencies are sufficient for the design of each filter). The NP in the Sinc convolution layer remains constant independently of L. Therefore, the SincNet provides a gain in a number of parameters defined by the equation (12):

$$Gain\ (\%) = \frac{NP_{CNN} - NP_{SincNet}}{NP_{CNN}} \times 100$$
$$= (1 - 2/L) \times 100 \qquad (12)$$

For example, the implementation of a SincNet architecture in the present BF diagnosis application provides a gain of 83.3% in terms of the number of parameters, compared to the conventional CNN.

4.5. Deep-SincNet application on the CWRU bearing database

In order to show the performance of the SincNet for a benchmark database, further tests based on the CWRU bearing database are achieved. For comparison reasons, this dataset is similar to that tested in [23], [24], [29]. The vibration signals are collected for the bearing faults IRF, ORF and Ball-F with holes of 0.007, 0.014, and 0.021 in. Data are acquired by

an accelerometer in the 6 o'clock position under different load levels (from 0 to 3 hp). The selected model is represented in table 5. It can be seen that the architecture is similar to that used for the first dataset, unless changing some hyper-parameters since the CWRU signals have different sampling parameters.

As shown in Fig.7, the t-SNE distribution of the final layer shows that the samples related to each class are clearly distinguished and separated. An accuracy of 99.95% ± 0.05 is reached, which proves the suitability of the SincNet-based models with different databases. For the diagnosis of the same bearing conditions, Gan et al. [23] combined the Wavelet Packet Transform with DBN in order to design a hierarchical network. Jia et al [24] used a Deep Neural Network (DNN) fed by Fourier features. In order to overcome the drawbacks of the Sparse Auto Encoders (SAE), Wang et al. [29] introduced a batch normalization technique into SAE-based DNN. In table 6, the accuracy reached by the Deep-SincNet is compared with those attained in [23], [24], [29], where the same bearing conditions were tested. The comparison with several potential approaches proves the competitive performances of the SincNet, which can be seen as an efficient alternative for fault diagnosis.

Table 5: Parameters of the model designed for the diagnosis of the CWRU dataset

Layer	Hyper parameters	Output Shape	Number of parameters
Input layer	-	(400, 1)	0
Sinc-convolutional layer	F = 30, L = 15	(386, 30)	100
Batch Normalization	-	(386, 30)	100
1-D MaxPooling	pooling size = 3	(128, 30)	0
Leaky Relu	α = 0.2	(128, 30)	0
1-D Convolutional Layer	F = 12 L = 5	(124, 12)	10010
Batch Normalization	-	(124, 12)	40
1-D Max Pooling	pooling size = 3	(41, 12)	0
Leaky ReLu	α = 0.2	(41, 12)	0
1-D Global Average Pooling	-	(12)	0
Fully connected layer with softmax	Number of units = 10	(10)	99

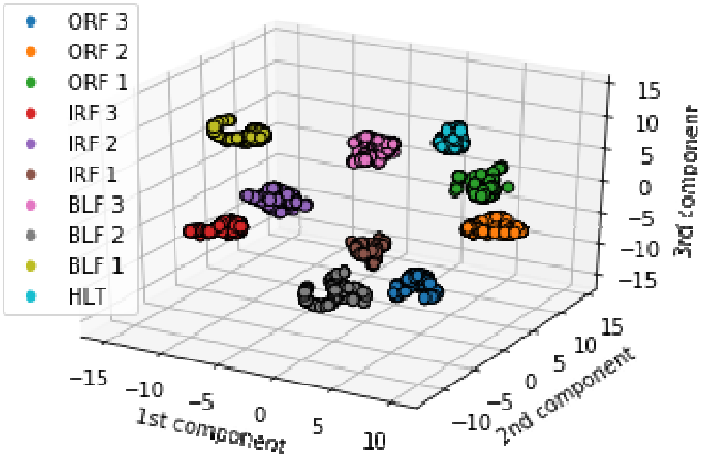


Fig. 7. t-SNE multi-class visualization of the CWRU dataset

5. Conclusion

This work proposed an end-to-end process for the diagnosis and the severity assessment of the bearing faults in IM. Various enhancements were achieved by the Deep-SincNet. First, the model automated the feature extraction step and extracted further fault descriptors as well as the conventional patterns related to the BF characteristic frequencies. Furthermore, the Deep-SincNet is characterized by its physical interpretability, as it uses explainable filters and learns purefrequency bandwidths. Accordingly, it can explore and recognize further fault patterns, generally unexploited by conventional techniques. The Deep-SincNet also offers a fast convergence, which is almost seven times faster than CNN with reduced computational cost. These advantages go with a high accuracy result in terms of classification of the different bearing faults and terms of the assessment of fault severity.

In summary, the proposed scheme is an intelligent end-to-end architecture, which uses raw vibration signals as input acquired with only one low-cost sensor as inputs and removes the need forfeature extraction and selection. Its efficiency was proven in terms of precision and sensitivity for two different databases under different working conditions and several levels of severity.

References

[1] Induction Motor Market: Global Industry Trends, Share, Size, Growth, Opportunity and Forecast 2018-2023, IMARC group, Dec. 2018.

[2] S.B. Johnson, T. J. Gormley et al., The Theory of System Health Management, in *System Health Management with Aerospace Applications*, Chichester, United Kingdom, John Wiley & Sons, 3-26, 2011.

[3] S. Khan and T. Yairi, A review on the application of deep learning in system health management, *Mech. Syst. Sig. Process*, 107, 241–265, 2018.

[4] F. B. Abid, S. Zgarni, and A. Braham, Distinct bearing faults detection in induction motor by a hybrid optimized SWPT and AINet-DAGSVM,*IEEE Trans. Energy Convers.*, 33 (4), 1692–1699, 2018.

[5] J. Harmouche, C. Delpha and D. Diallo, A global approach for the classification of bearing faults conditions using spectral features, IECON 2013, Vienna, 7352-7357, 2013.

[6] R. Abdelkader, A. Kaddour, A. Bendiabdellah and Z. Derouiche, Rolling Bearing Fault Diagnosis Based on an Improved Denoising Method Using the Complete Ensemble Empirical Mode Decomposition and the Optimized Thresholding Operation, in *IEEE Sensors Journal*, 18 (6), 7166-7172, 2018.

[7] Z. Mezni, C. Delpha, D. Dialloand A. Braham,Performance of Bearing Ball Defect Classification Based on the Fusion of Selected Statistical Features. *Entropy*, 24(9), 1251, 2022.

[8] R.C. Eddine and B. Slimane, Detection of bearing defects using Hilbert envelope analysis and fast kurtogram demodulation method. *Journal of Electrical Systems*, 16, 2020.

[9] KC Kompella, M. Deekshit, Venu Gopala Rao, and R. Srinivasa Rao. Bearing fault diagnosis in 3 phase induction machine using current spectral subtraction with different wavelet transform techniques, *Journal of Electrical Systems*, 13 (1),143-159, 2017.

[10] M. Altaf, T. Akram, M. A. Khan, M. Iqbal,M.M.I. Chand C.H.Hsu, A New Statistical Features Based Approach for Bearing Fault Diagnosis Using Vibration Signals, *Sensors*, 22(5), 2012, 2022.

[11] M. Hosseinpour-Zarnaq, M. Omid, and E. Biabani-Aghdam. Fault diagnosis of tractor auxiliary gearbox using vibration analysis and random forest classifier, *Information Processing in Agriculture*, 2021.

[12] He, Cheng, T. Wu, R. Gu, Z. Jin, R. Ma, and H. Qu. , Rolling bearing fault diagnosis based on composite multiscale permutation entropy and reverse cognitive fruit fly optimization algorithm–Extreme learning machine, *Measurement*, 173, 108636, 2021.

[13] Q. Hu, A. Qin, Q. Zhang, J. He and G. Sun, Fault Diagnosis Based on Weighted Extreme Learning Machine With Wavelet Packet Decomposition and KPCA, in *IEEE Sensors Journal*, 18 (20), 8472-8483, 2018.

[14] J. Xiancheng, Y. Ren, H. Tang, and J. Xiang, DSmT-based three-layer method using multi-classifier to detect faults in hydraulic systems, *Mechanical Systems and Signal Processing,* 153, 107513, 2021.

[15] Y. Orhan, An automated faults classification method based on binary pattern and neighborhood component analysis using induction motor, *Measurement*, 168, 108323, 2021.

[16] Y. Hong, M. Kim, H. Lee, J. J. Park and D. Lee, Early Fault Diagnosis and Classification of Ball Bearing Using Enhanced Kurtogram and Gaussian Mixture Model, in *IEEE Trans. on Instrum. and Meas*. 68, 4746-4755, 2021.

[17] 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. Sig. Process*, 42 (1-2), 314–334, 2014.
[18] A. Khadersab, Machine learning algorithms for rotating machinery bearing fault diagnostics, *Materials Today: Proceedings*, 2021.
[19] MM. Islam and J.M Kim, Reliable multiple combined fault diagnosis of bearings using heterogeneous feature models and multiclass support vector Machines, *Reliability Engineering & System Safety*, 184, 55-66, 2019.
[20] C. Scheffer and P. Girdhar, Practical machinery vibration analysis and predictive maintenance. *Elsevier*, 2004.
[21] J. Harmouche, C. Delpha, and D. Diallo, Improved fault diagnosis of ball bearings based on the global spectrum of vibration signals,*IEEE Trans. Ener. Conv.*, 30 (1), 376–383, 2015.
[22] H. Shao, H. Jiang, X. Li, and T. Liang, Rolling bearing fault detection using continuous deep belief network with locally linear embedding, *Computers in Industry*, 96, 27–39, 2018.
[23] M. Gan, C. Wang et al. , Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings, *Mech. Syst.and Sig. Process*., 72, 92-104, 2016.
[24] F. Jia, Y. Lei, J. Lin et al., Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data, *Mech. Syst.and Sig. Process*, 72, 303-315, 2016.
[25] H. Oh, J. H. Jung, B. C. Jeon, and B. D. Youn, Scalable and unsupervised feature engineering using vibration-imaging and deep learning for rotor system diagnosis, in *IEEE Trans. Ind. Electron.,* 65 (4), 3539–3549, 2018.
[26] S. Shao, S. McAleer, R. Yan and P. Baldi, Highly Accurate Machine Fault Diagnosis Using Deep Transfer Learning,in *IEEE Transactions on Ind. Inform.*, 15 (4), 2446-2455, 2019.
[27] L. Wen, X. Li, L. Gao, and Y. Zhang, A new convolutional neural network-based data-driven fault diagnosis method, in *IEEE Trans. Ind. Electron*., 65 (7), 5990–5998, 2018.
[28] H. O. A. Ahmed, ML Wong, and A. Nandi, Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features, *Mech. Syst.and Sig. Process*, 99, 459-477, 2018.
[29] W. Mao, W. Feng, Y. Liu, D. Zhang, and X. Liang, A new deep auto-encoder method with fusing discriminant information for bearing fault diagnosis, *Mechanical Systems and Signal Processing*, 150, 107233, 2021.
[30] J. Yang, B. Weimin, L. Yanming, L. Xiaoping, W. Junjie, N. Yue, and L. Jin, Joint pairwise graph embedded sparse deep belief network for fault diagnosis, *Engineering Applications of Artificial Intelligence*, 99, 104149, 2021.
[31] Ji. Mengyu, G. Peng, J. He, S. Liu, Z. Chen, and S. Li., A Two-Stage, Intelligent Bearing-Fault-Diagnosis Method Using Order-Tracking and a One-Dimensional Convolutional Neural Network with Variable Speeds, *Sensors,* 21, (3), 675, 2021.
[32] W. Lu, Y. Li, Y. Cheng et al., Early fault detection approach with deep architectures,*IEEE Trans. on Instrum. and Meas.*, 67 (7), 1679-1689, 2018.
[33] M. Ravanelli and Y. Bengio, Speaker recognition from raw waveform with sincnet, In 2018 *IEEE Spoken Language Technology Workshop* (SLT), 1021-1028, 2018.
[34] F. Ben Abid, M. Sallem and A. Braham, Robust Interpretable Deep Learning for Intelligent Fault Diagnosis of Induction Motors,in *IEEE Transactions on Instrumentation and Measurement*, 69 (6), 3506-3515, 2020.