



Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals



Chuan Li ^{a,b,*}, René-Vinicio Sanchez ^b, Grover Zurita ^b, Mariela Cerrada ^b,
Diego Cabrera ^b, Rafael E. Vásquez ^c

^a Research Center of System Health Maintenance, Chongqing Technology and Business University, Chongqing 400067, China

^b Department of Mechanical Engineering, Universidad Politécnica Salesiana, Cuenca, Ecuador

^c Department of Mechanical Engineering, Universidad Pontificia Bolivariana, Medellín, Colombia

ARTICLE INFO

Article history:

Received 14 July 2015

Received in revised form

29 January 2016

Accepted 3 February 2016

Available online 15 February 2016

Keywords:

Deep learning

Data fusion

Gearbox

Acoustic emission

Vibration signal

ABSTRACT

Fault diagnosis is an effective tool to guarantee safe operations in gearboxes. Acoustic and vibratory measurements in such mechanical devices are all sensitive to the existence of faults. This work addresses the use of a deep random forest fusion (DRFF) technique to improve fault diagnosis performance for gearboxes by using measurements of an acoustic emission (AE) sensor and an accelerometer that are used for monitoring the gearbox condition simultaneously. The statistical parameters of the wavelet packet transform (WPT) are first produced from the AE signal and the vibratory signal, respectively. Two deep Boltzmann machines (DBMs) are then developed for deep representations of the WPT statistical parameters. A random forest is finally suggested to fuse the outputs of the two DBMs as the integrated DRFF model. The proposed DRFF technique is evaluated using gearbox fault diagnosis experiments under different operational conditions, and achieves 97.68% of the classification rate for 11 different condition patterns. Compared to other peer algorithms, the addressed method exhibits the best performance. The results indicate that the deep learning fusion of acoustic and vibratory signals may improve fault diagnosis capabilities for gearboxes.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Gearboxes have been widely used as key machine components for delivering torque and providing speed conversions from some rotating power sources to other devices. Any failures in gearboxes may introduce unwanted downtime, expensive repair procedures and even human casualties. As an effective component for condition-based maintenance [1], the fault diagnosis has gained much attention in order to guarantee safe operations of gearboxes.

Gearbox conditions can be reflected by measurements of vibratory [2], acoustic [3], thermal [4], electrical [5], and oil-based signals [6]. In vibration-based gearbox fault diagnostics, Cheng et al. [7] proposed an order tracking technique using local mean decomposition of the vibration signal for gear fault diagnosis. Wang et al. [8] used wavelet decomposition for robust health evaluation of gearbox subject to tooth failure. Lei et al. [9] analyzed vibration characteristics in both time and frequency domains for the diagnostics of the planetary gearboxes. In addition to vibratory signals, acoustic ones are also

* Corresponding author at: Research Center of System Health Maintenance, Chongqing Technology and Business University, Chongqing 400067, China.
Tel.: +86 23 6276 8469.

E-mail address: chuanli@21cn.com (C. Li).

sensitive to the existence of gearbox faults. Li et al. [10] applied adaptive morphological gradient lifting wavelet to detect gear fault. Chacon et al. [11] used acoustic emission (AE) to detect shafts angular misalignments. Hamel et al. [12] researched the influence of the oil film thickness on helical gearboxes fault detection via the AE signal. Abad et al. [13] applied acoustic signals for gearboxes fault detection using the discrete wavelet transform and artificial neural networks.

To explore different symptoms for gearboxes fault diagnosis, Rafael et al. [14] combined acoustic emissions and vibration measurements to detect gearbox failures. Li et al. [15] used fault features from vibration and acoustic emission signals for gearboxes fault detection using the K-nearest neighbor (KNN) algorithm. The combined vibratory and AE signature was suggested by Soua et al. [16] as a pre-requisite for condition monitoring of a wind turbine gearbox. Khazaee et al. [17] developed a fault classifier using data fusion of the vibration and the acoustic signals for planetary gearboxes using the Dempster–Shafer evidence theory.

Usually, gearboxes are operated within harsh environments with heavy noises and interferences. For better diagnosis performance, therefore, the fault-sensitive features should be extracted from raw signals. An optimal mathematical morphology demodulation technique was reported by Li and Liang [18] to extract the impulsive feature for bearing defect diagnosis. Raad et al. [19] employed the cyclostationarity as an indicator to the diagnostics of the gears. A criterion fusion approach was reported by Li et al. [20] to optimal demodulation of vibration signals. Chen et al. [21] proposed an intelligent diagnosis model including wavelet support vector machines and immune genetic algorithms for gearboxes. A generalized synchrosqueezing transform was developed to diagnose gearbox faults and bearing defects [22,23]. To reduce the dimensionality of data, Hinton and Salakhutdinov [24] developed a deep learning framework, and further applied to feature extraction [25], classification [26], regression [27] for images, signals, and time series. Under deep learning framework, deep belief networks were introduced to diagnose reciprocating compressor valves [28] and gearboxes [29]. A deep belief learning based health state classification was applied for different dataset such as iris dataset and wine dataset [30]. The researches indicate that deep learning is capable of better extracting features compared to traditional learning approaches.

The existing reports directly applied deep learning for the fault diagnosis using one signal. In this work we propose the fusion of acoustic and vibratory signals as a fault diagnostics tool for gearboxes using a deep random forest fusion (DRFF) technique. The fault-sensitive features of such signals are extracted using the statistical parameters of the wavelet packet transform (WPT) and deep learning with deep Boltzmann machines (DBMs). Instead of a traditional combination of the two signals, we suggest random forest (RF) as a data fusion tool for the integration of the deep feature representations. With deep learning feature representations and data fusion strategies, this work improves the performance of gearbox fault diagnosis, which is validated by experiments and comparisons with peer techniques.

The structure of the paper is presented as follows. Deep representations of the AE and the vibratory features using WPT and DBM are introduced in Section 2. In Section 3, the RF is proposed to fuse deep representations of the AE and the vibratory features in an integrated fashion; the application of the proposed DRFF method to the gearbox fault diagnosis is also detailed in this section. In Section 4, the gearbox fault diagnosis experiments are carried out to evaluate the present approach. Finally, conclusions are given in Section 5.

2. Deep learning for the condition feature representations

2.1. WPT statistical parameters of the gearbox condition measurements

The operating condition of a gearbox can be determined using measurements from different sensors. In this research, both the AE sensor and the accelerometer are defined as the gearbox measurements $x(t)$ as

$$x(t) = \{x^{(1)}(t), x^{(2)}(t)\}, \quad (1)$$

where $x^{(1)}(t)$ and $x^{(2)}(t)$ denote the acoustic and the vibratory signals, respectively. The gearbox condition measurements can be decomposed into different depths using the WPT as [31]

$$Wx_n(j, k) = \langle x(t), \mu_n(t, j, k) \rangle = \int_{-\infty}^{+\infty} x(t) \mu_n(t, j, k) dt. \quad (2)$$

where $\langle *, * \rangle$ denotes the inner product operator, μ_n the n -th wavelet packet function ($n=0,1,2,\dots,2^j$), j the level and k the wavelet coefficient. The gearbox condition can be reflected by the information included in different wavelet packet nodes. Usually, statistical parameters are good indicators for extracting the condition information. In this research, the following statistical parameters for each node are used [32]

$$\begin{aligned} S_1(j, n) &= \frac{\max |Wx_n(j, k)|}{\sqrt{\frac{1}{K} \sum_{k=1}^K (Wx_n(j, k))^2}}, S_2(j, n) = \frac{\sqrt{\frac{1}{K} \sum_{k=1}^K (Wx_n(j, k))^2}}{\frac{1}{K} \sum_{k=1}^K |Wx_n(j, k)|}, S_3(j, n) = \frac{1}{K} \sum_{k=1}^K |Wx_n(j, k)|, \\ S_4(j, n) &= \left(\frac{1}{K} \sum_{k=1}^K \sqrt{|Wx_n(j, k)|} \right)^2, S_5(j, n) = \frac{1}{K} \sum_{k=1}^K (Wx_n(j, k))^4, S_6(j, n) = \frac{1}{K} (Wx_n(j, k))^2, \end{aligned}$$

$$S_7(j, n) = \frac{\max |Wx_n(j, k)|}{\left(\frac{1}{K} \sum_{k=1}^K \sqrt{|Wx_n(j, k)|}\right)^2}, S_8(j, n) = \frac{\max |Wx_n(j, k)|}{\frac{1}{K} \sum_{k=1}^K |Wx_n(j, k)|}, \text{ and } S_9(j, n) = \frac{1}{K} (Wx_n(j, k))^3. \quad (3)$$

where K stands for the length of the wavelet coefficients of $WP(j, n)$, and S_1, S_2, \dots, S_9 represent the crest factor, shape factor, absolute mean amplitude, square root amplitude, kurtosis, variance values, clearance factor, impulse indicator, and skewness factor of the wavelet coefficients of $WP(j, n)$, respectively.

With the maximum decomposition level J ($j=1, 2, \dots, J$), the WPT statistical parameter set Sx of the gearbox measurement x (t) can be extracted as

$$x(t) \leftrightarrow Sx(J) = [S_1(1, 1), \dots, S_9(1, 1), S_1(1, 2), \dots, S_9(1, 2), \dots, S_1(J, 2^J), \dots, S_9(J, 2^J)] \quad (4)$$

where $Sx \in R^{9 \times (2^{J+1} - 2)}$. Combining (1) and (7) yields

$$x = \{x^{(1)}, x^{(2)}\} \leftrightarrow Sx = \{Sx^{(1)}, Sx^{(2)}\}. \quad (5)$$

In (8) the acoustic and the vibratory signals x of the gearbox can be represented using the WPT statistical parameter set Sx . In the following subsection, Sx are exploited by deep learning to generate the gearbox condition features.

2.2. Deep learning of the condition features using the DBM

In the deep learning techniques, many layers of nonlinear information processing are used for feature extraction and transformation, as well as for pattern analysis and classification [33]. During the past several years, the deep learning has made important influences on a wide range of applications such as image processing, computer vision, phonetic recognition, voice search, semantic utterance classification, and hand-writing recognition [34].

Until recently, most of feature representation techniques used “shallow” structured architectures, which typically contain one or two layers of nonlinear feature transformations. For example, support vector machines (SVMs) exploit a linear pattern separation model with one or zero feature transformation layer [35]. Through being named as artificial intelligence, the shallow learning techniques differ from the human information processing mechanisms. Typically, human beings extract complex structure and build internal representation from rich sensory inputs with very deep architectures. This encouraged researchers to develop effective deep learning algorithms. In this paper, the DBM is exploited for the feature representation of the gearbox conditions. In other words, for each signal (AE or vibratory), a DBM structure is used for the feature representation.

The DBM can be constructed using multiple restricted Boltzmann machines (RBMs). A RBM is a network of symmetrically coupled stochastic binary units. A set of visible units $\mathbf{v} \in [0, 1]$ and a set of hidden units $\mathbf{h} \in [0, 1]$ as well as the connections restricted between the visible and the hidden neurons contribute to the RBM. By stacking several RBMs, one can form the hierarchical structure of the DBM which contains a single visible layer \mathbf{v} and L hidden layers $\mathbf{h}^{(1)}, \dots, \mathbf{h}^{(l)}, \dots, \mathbf{h}^{(L)}$.

Typical DBMs process only binary input data, which limits considerably their application area. To accommodate real-valued input data, a popular used method is to replace the binary visible neurons with Gaussian ones [36]. With this modification, the energy for a state $[\mathbf{v}^T, \mathbf{h}^{(1)T}, \dots, \mathbf{h}^{(L)T}]^T$ of the DBM is defined as

$$E(\mathbf{v}, \mathbf{h}^{(1)}, \dots, \mathbf{h}^{(L)} | \theta) = \sum_{i=1}^{N_v} 2(v_i - b_i)^2 / \sigma_i^2 - \sum_{i=1}^{N_v} \sum_{j=1}^{N_1} w_{ij} v_i h_j^{(1)} / \sigma_i^2 - \sum_{l=1}^L \sum_{j=1}^{N_l} b_j^{(l)} h_j^{(l)} - \sum_{l=1}^{L-1} \sum_{j=1}^{N_l} \sum_{k=1}^{N_{l+1}} w_{jk}^{(l)} h_j^{(l)} h_k^{(l+1)}, \quad (6)$$

where $\theta = \{\mathbf{w}, \mathbf{b}\}$ stands for the model parameters, $w_{ij} \in \mathbf{w}$ is the weight of the synaptic connection between the i -th visible neuron and the j -th hidden neuron, $b_i \in \mathbf{b}$ represents the i -th bias term, N_v is the length of visible layer, N_l is the length of the L -th hidden layer, and σ_i denotes the standard deviation of the conditional probability of the i -th visible neuron, which is given by

$$p(v_i | \mathbf{h}^{(1)}, \theta) = N\left(v_i | \sum_{j=1}^{N_1} h_j^{(1)} w_{ij} + b_i, \sigma_i^2\right), \quad (7)$$

where $N(\cdot)$ is the probability density function of the normal distribution. Similarly, the conditional probabilities of the hidden neurons can be derived as

$$p(h_j^{(l)} | \mathbf{h}^{(l-1)}, \mathbf{h}^{(l+1)}, \theta) = M\left(\sum_{i=1}^{N_{l-1}} h_i^{(l-1)} w_{ij}^{(l-1)} + \sum_{k=1}^{N_{l+1}} h_k^{(l+1)} w_{jk}^{(l)} + b_j^{(l)}\right), \quad l \neq 1, L, \quad (8)$$

where $M(\cdot)$ is the sigmoid function. Eq. (11) is not applicable for the first hidden layer ($l=1$) and the last hidden layer ($l=L$). For the last hidden layer one should set $N_{L+1}=0$; for the first hidden layer, one has $\mathbf{h}^{(l-1)} = \mathbf{v}$. Hence, (11) should be rewritten for the two special cases as

$$p(h_j^{(1)} | \mathbf{v}, \mathbf{h}^{(2)}, \theta) = M\left(\sum_{i=1}^{N_v} v_i w_{ij} / \sigma_i^2 + \sum_{k=1}^{N_2} h_k^{(2)} w_{jk}^{(1)} + b_j^{(1)}\right), \quad l=1, \quad (9)$$

$$p(h_j^{(L)} | \mathbf{h}^{(L-1)}, \theta) = M \left(\sum_{i=1}^{N_{L-1}} h_i^{(L-1)} w_{ij}^{(L-1)} + b_j^{(L)} \right), l = L. \quad (10)$$

When being applied for the classification tasks, the typical DBM uses a softmax layer or 1-of- I encoding at the top. To specify the only right class \hat{i} from I possible candidates (namely, $i = 1, 2, \dots, I$), the i -th input for the softmax layer is expressed as

$$A_i = \sum_{j=1}^{N_{L-1}} h_j^{(L-1)} w_{ij}^{(L-1)}, \quad (11)$$

The softmax layer should have I nodes with discrete probabilities, i.e., P_1, P_2, \dots, P_I as given by

$$P_i = \frac{\exp(A_i)}{\sum_{i=1}^I \exp(A_i)}, \text{ and } \sum_{i=1}^I P_i = 1. \quad (12)$$

By feeding $Sx^{(1)}$ and $Sx^{(2)}$ into two DBMs, i.e., $\text{DBM}^{(1)}$ and $\text{DBM}^{(2)}$, respectively, the gearbox condition features of the acoustic and the vibratory signals are represented by

$$x^{(1)}(t) \leftrightarrow P_i^{(1)}, \text{ and } x^{(2)}(t) \leftrightarrow P_i^{(2)}. \quad (13)$$

3. RF fusion of deep representation features for the gearbox fault diagnosis

To simultaneously use both, acoustic and vibratory features, the RF is introduced to fuse them into an integrated aspect. The proposed DRFF method is then applied to the gearbox fault diagnosis.

3.1. Deep features fusion using the RF

With the feature representation given by (16), gearbox measurements using either acoustic or vibratory signals would result in a healthy condition class \hat{i} , which can be determined through a voting classification mechanism from (15) as [37]

$$\hat{i}^{(1)} = \arg \max_i P_i^{(1)} = \arg \max_i A_i^{(1)}, \text{ and } \hat{i}^{(2)} = \arg \max_i P_i^{(2)} = \arg \max_i A_i^{(2)}. \quad (14)$$

Unfortunately, the two measurements may cause two class results. If $\hat{i}^{(1)} = \hat{i}^{(2)}$, the gearbox condition pattern will be clear. On the other hand, if $\hat{i}^{(1)} \neq \hat{i}^{(2)}$, one has to determine which one (or, sometimes none of the ones) is the right result. To tackle this problem, we suggest the RFs as data fusion tools to replace the voting classification mechanism.

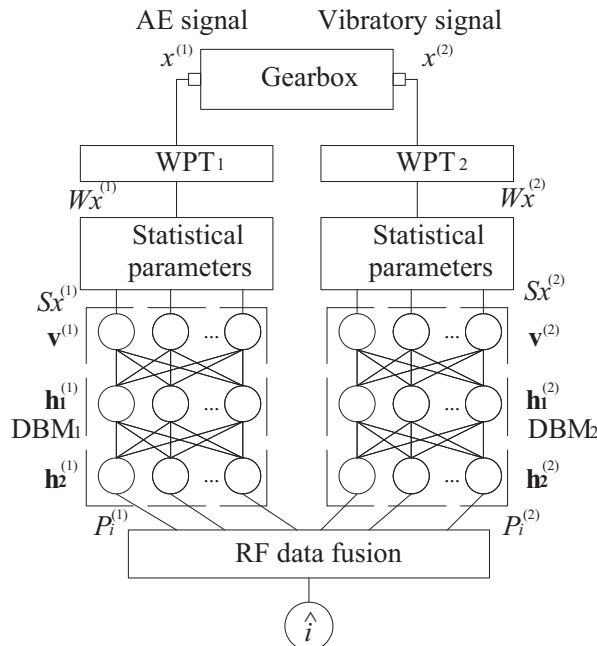


Fig. 1. Schematic structure of the proposed DRFF model for the gearbox fault diagnosis.

The RFs are an ensemble machine learning method for classification, regression, among others. Instead of using direct voting classification approach towards $\hat{i}^{(1)}$ and $\hat{i}^{(2)}$, the deep feature representations of $p_i^{(1)}$ and $p_i^{(2)}$ are employed as the input of the RF data fusion. The input of the RF is then expressed as

$$P = [P_1^{(1)}, \dots, P_I^{(1)}, P_1^{(2)}, \dots, P_I^{(2)}]. \quad (15)$$

There are B trees as an ensemble $\{T_1(P), \dots, T_b(P), \dots, T_B(P)\}$ to compose the RF. With the input variable P , the ensemble produces B outputs $\{T_1(P), \dots, T_b(P), \dots, T_B(P)\}$ as the prediction results for all the trees. The outputs of all the trees are aggregated to produce the final healthy condition class \hat{i} , which is the class predicted by the majority of the trees, as illustrated by (17). The RF algorithm for the classification is summarized as [38]: (1) for $b=1$ to B , choose a bootstrap sample from the training data set; (2) for each bootstrap, grow a tree T_b by recursively applying the steps to achieve the maximum tree size and not pruned back; (3) generate the outputs the trees of the RF as $Rf = \{\hat{i}_1, \dots, \hat{i}_I\}$; and (4) determine the final class corresponding to the input data as

$$\hat{i} = \arg \max_i \{\hat{i}_1, \dots, \hat{i}_I\}. \quad (16)$$

3.2. Overview of the present DRFF approach for the gearbox fault diagnosis

Through the deep learning of the WPT statistical parameters from acoustic and vibratory signals, the RF can be used to fuse the deep features into an integrated aspect. Then, we suggest the use 1 visible layer and 2 hidden layers for each DBM. With these constituents, one can develop the deep random forest fusion (DRFF) structure for the gearbox diagnosis as shown in Fig. 1.

After getting the DRDF model, it is necessary to train the model using the training dataset. DBM₁, DBM₂, and RF can be trained separately. The output results $p_i^{(1)}$ and $p_i^{(2)}$ of the trained DBMs should be fed into the RF for the data fusion training. The details of training the DBMs and the RFs can be found in [39]. Having introduced the procedures separately, the application of the present DRDF technique for the gearbox fault diagnosis can be summarized as follows:

- Step 1. Collect the measurements $x^{(1)}(t)$ and $x^{(2)}(t)$ from the AE sensor and the vibratory accelerometer, and define the healthy condition patterns for the gearbox to be diagnosed.
- Step 2. Perform the WPT for $x^{(1)}(t)$ and $x^{(2)}(t)$ to respectively obtain $Wx_n^{(1)}(j,k)$ and $Wx_n^{(2)}(j,k)$ using Eq. (5).
- Step 3. Calculate the WPT statistical parameter sets $Sx^{(1)}$ and $Sx^{(2)}$ based on (6)–(8).
- Step 4. Construct DBM₁, DBM₂ and employ $Sx^{(1)}$ and $Sx^{(2)}$ as their input layers, respectively.
- Step 5. At the output of DBM₁ and DBM₂, construct an RF data fusion using the input given by (18).
- Step 6. Train the DBM₁ and DBM₂ separately, and feed the results of the two DBMs into the RF to complete the DRFF model training.
- Step 7. When applying the trained DRFF model to the gearbox diagnose, input new acoustic and vibratory signals into the trained DRFF model to obtain the diagnosis result.

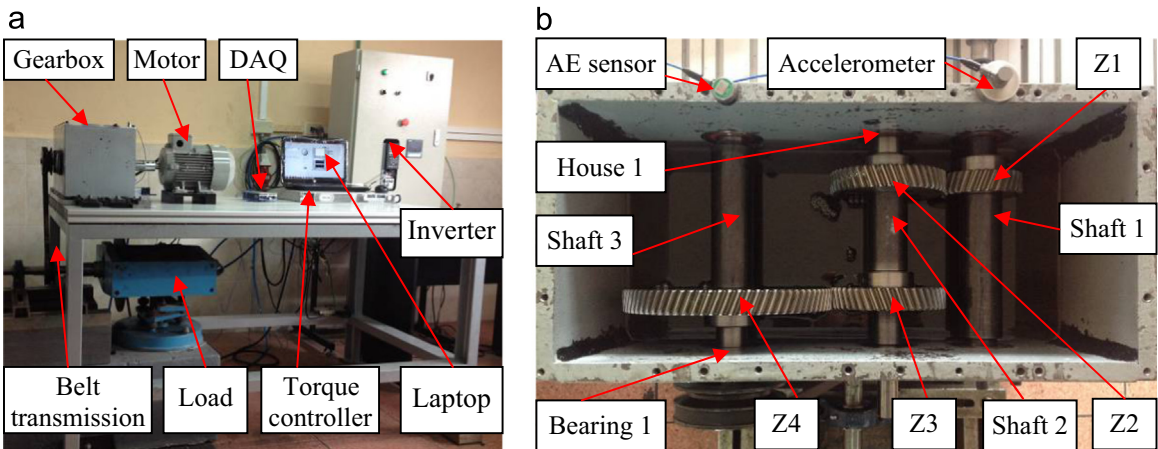


Fig. 2. Experimental set-up: (a) the picture of the set-up; and (b) the internal configuration of the gearbox.

4. Experiments and discussions











A gearbox experimental set-up is constructed to collect AE and vibratory measurements. The present DRFF technique is then applied for the gearbox fault diagnosis experiments. Different peer approaches are also introduced for the comparisons with the proposed technique.

4.1. Experimental set-up

To validate the effectiveness of the proposed method, we carried out experiments on a gearbox fault diagnosis experimental platform (fabricated by the laboratory of the Universidad Politécnica Salesiana, Ecuador). As shown in Fig. 2(a), the gearbox to be diagnosed is driven by a motor (SIEMENS, 3~, 2.0 HP) through a coupling. The speed of the motor is controlled by a frequency inverter (DANFOSS VLT 1.5 kW). The output shaft of the gearbox is connected to an electromagnetic torque load (ROSATI, maximum 8.83 kW) via a belt transmission. The electromagnetic torque load was controlled by a torque controller (TDK-Lambda, GEN 100-15-IS510), which allows one to adjust the torque of the load manually. The acoustic signals and vibration signals of the gearbox are collected by an AE sensor (PANAMETRICS V109) and an accelerometer (PCB ICP 353C03), respectively. The outputs of the AE sensor and the accelerometer are fed into a laptop (HP Pavilion g4-2055la) through a data acquisition box (DAQ, NI cDAQ-9234).

Fig. 2(b) shows the internal configuration of the gearbox. There are 3 shafts and 4 gears composing a two-stage transmission of the gearbox. An input helical gear ($Z_1=30$, modulus=2.25, impact angle 20° and helical angle 20°) is installed on the input shaft (shaft 1). Two intermediate helical gears ($Z_2=Z_3=45$) are installed on an intermediate shaft (shaft 2) for the transmission between the input gear and the output gear ($Z_4=80$, installed on the output shaft, i.e., shaft 3). The faulty components used in the experiments included gears Z_1 , Z_2 , Z_3 , Z_4 , bearing 1 and bearing house 1 as shown in Fig. 2(b). With the aforementioned faulty components, different faults patterns as shown in Table 1 were set for the gearbox diagnosis experiments.

Table 1
Condition patterns of the gearbox configuration.

Faulty pattern	Faulty component	Faulty detail	Input speed	Load	Faulty photo
A	N/A	N/A	300, 600, 900 rpm	Zero, small, great	N/A
B	Gear Z_1	Worn tooth	300, 600, 900 rpm	Zero, small, great	
C	Gear Z_2	Chaffing tooth	300, 600, 900 rpm	Zero, small, great	
D	Gear Z_3	Pitting tooth	300, 600, 900 rpm	Zero, small, great	
E	Gear Z_3	Worn tooth	300, 600, 900 rpm	Zero, small, great	
F	Gear Z_4	Chipped tooth	300, 600, 900 rpm	Zero, small, great	
G	Gear Z_4	Root crack tooth	300, 600, 900 rpm	Zero, small, great	
H	Bearing 1	Inner race fault	300, 600, 900 rpm	Zero, small, great	
I	Bearing 1	Outer race fault	300, 600, 900 rpm	Zero, small, great	
J	Bearing 1	Ball fault	300, 600, 900 rpm	Zero, small, great	
K	House 1	Eccentric	300, 600, 900 rpm	Zero, small, great	

4.2. Experimental results

With the experimental configuration shown in Table 1, 11 patterns with 3 different load conditions and 3 different input speeds were applied during the experiments. For each pattern, load, and speed condition, the tests were repeated 5 times. In each experiment, AE and vibratory signals were collected with 24 durations each covering 0.4096 s. The sampling frequency for AE and vibratory signals were set at 50 kHz and 10 kHz, respectively. In this way, 11,880 AE signals (i.e., $[x_1^{(1)}(t), \dots, x_{11,880}^{(1)}(t)]$) and 11,880 vibratory signals (i.e., $[x_1^{(2)}(t), \dots, x_{11,880}^{(2)}(t)]$) corresponding to 11 condition patterns (i.e., [A, ..., K]) were collected from the gearbox fault diagnosis experiments.

The DRFF model was developed with the following parameters: mother wavelet for the WPT₁ and WPT₂ was Daubechies wavelet of order 5 (i.e., db5); decomposition level for WPT₁=5; decomposition level for WPT₂=4; number of the neurons for the hidden layer 1 of DBM₁=300; number of the neurons for the hidden layer 2 of DBM₁=300; number of the pre-training epochs of DBM₁=50, number of the fine-tuning epochs of DBM₁=150; number of neurons for the hidden layer 1 of DBM₂=200; number of neurons for the hidden layer 2 of DBM₂=200; number of the pre-training epochs of DBM₁=50, number of fine-tuning epochs of DBM₁=100; and the number of trees of the RF=50. When applying DBM for deep learning, network parameters should be carefully chosen. The number of the neurons is an important factor influencing feature extraction and data fusion results. Usually a small number of the neurons introduces inferior learning result, with less computational burden. On the contrary, more neurons often generate better learning result, at the cost of computational resources. Hence in the experiments we empirically chose the number as 300, which is a compromise between the computation burden and the learning performance. The visible layer was employed as the input layer of the deep learning networks. Hence only 1 visible layer was used for each DBM. As for the number of the hidden layers, theoretically more layers may generate better classification ratio. However, this will introduce much computational burden. We used 2 hidden layers in this work where good results were obtained.

To train the DRFF model, we random chose 7920 AE and vibratory signals as the inputs. With the trained DRFF model, the rest 3960 AE and vibratory signals were exploited for the test of the fault diagnosis performances. Among all the 3960 tests, the trained DRFF generated 3868 correct classifications. In other words, the present DRFF produced 97.68% classification rate for the gearbox fault diagnosis experiments. In addition to the total classification rate, Fig. 3 shows the classification rates of the DEFF model for different faulty patterns.

To detail the fault diagnosis of the present DRFF model, let's take 9465 test as an example. This test was performed with the following conditions: fault pattern J (ball fault of bearing 1), input speed 600 rpm, and small load. Fig. 4(a) and (b) shows raw signals of the AE sensor and the accelerometer (i.e., $x_{9465}^{(1)}(t)$ and $x_{9465}^{(2)}(t)$), respectively. Their WPT statistical parameters $Sx_{9465}^{(1)}$ and $Sx_{9465}^{(2)}$ are displayed in Fig. 4(c) and (d), respectively. Fig. 4(e) and (f) shows their deep learning representations $Pt_{9465}^{(1)}$ and $Pt_{9465}^{(2)}$ using DBM₁ and DBM₂, respectively. Comparing Fig. 4(e) with (f) indicates that different classes are generated by the AE sensor and the accelerometer for the same gearbox condition. With the direct voting classification mechanism of the typical DBM, the AE measurement diagnoses the gearbox's faulty pattern K (eccentric bearing house 1), while the vibratory measurement judges the gearbox's fault as pattern I (outer race fault of bearing 1). Unfortunately, both the AE and the vibratory diagnoses are all wrong by themselves. Fig. 4(g) plots the outputs of the RF trees. It is shown that the present DRFF model diagnoses the current condition as pattern J. This reveals that, even if both the AE and the vibratory measurements result in wrong patterns, the proposed DRFF model with data fusion may lead to right classification results.

4.3. Comparisons and discussions

To further validate the effectiveness of the proposed DRFF technique, the same data were used for the fault diagnosis with some peer methods. The presented method is compared to three different techniques (1) deep learning with other data fusion strategies; (2) deep learning without data fusion; and (3) shallow learning methods.

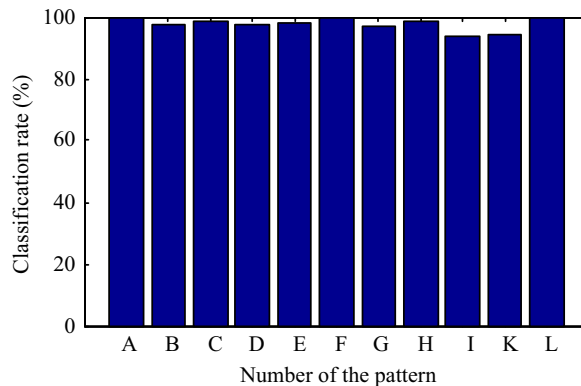


Fig. 3. Classification rates for different fault patterns using the test dataset.

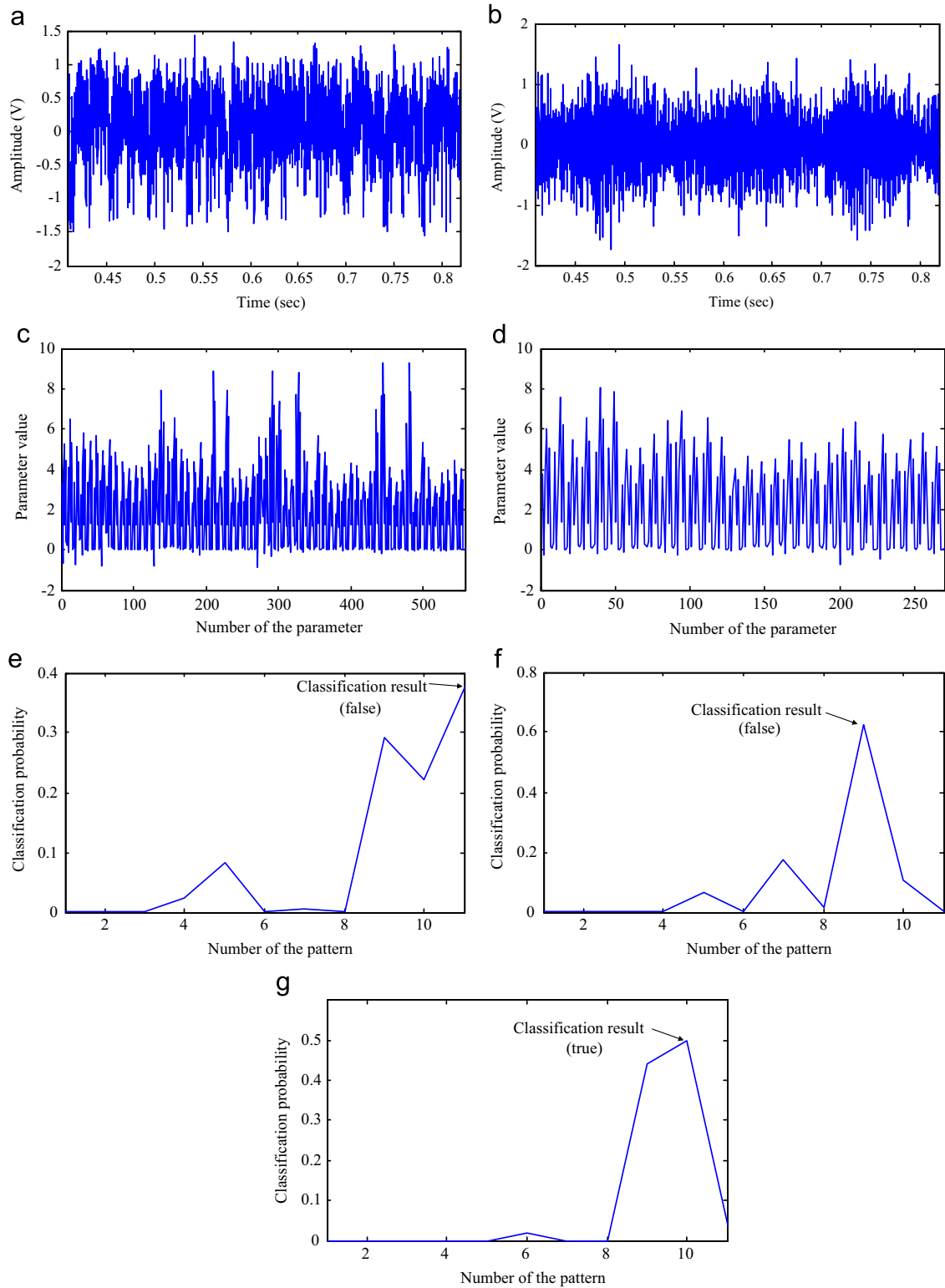


Fig. 4. Fault diagnosis detail for the 9465 test: (a) raw signal of the AE measurement; (b) raw signal of the vibratory measurement; (c) WPT statistical parameters of the AE signal; (d) WPT statistical parameters of the vibratory signal; (e) deep learning representation of the DBM₁; (f) deep learning representation of the DBM₂; and (g) RF outputs.

Table 2

Comparisons with peer classification methods.

Category	Method	Classification rate (%)
Deep learning with the RF	Present DRFF	97.68
Deep learning with other data fusion tools	Replace RF with SVM for DRFF	96.45
	Replace RF with KNN for DRFF	95.58
Deep learning without the data fusion	AE signal for DBM	79.77
	AE signal for DRFF	83.94
	Vibratory signal for DBM	88.74
	Vibratory signal for DRFF	89.49
Shallow learning	AE signal for SVM	58.94
	AE signal for KNN	32.42
	AE signal for RF	87.20
	Vibratory signal for SVM	71.57
	Vibratory signal for KNN	51.84
	Vibratory signal for RF	88.43

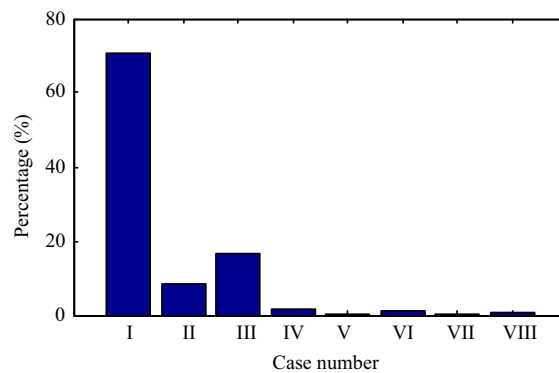


Fig. 5. Statistics of the classification results using the DRFF, DBM₁ and DBM₂, where case I – valid DRFF, DBM₁ and DBM₂; case II – valid DRFF and DBM₁; case III – valid DRFF and DBM₂; case IV – valid DRFF; case V – valid DBM₁ and DBM₂; case VI – valid DBM₂; case VII – valid DBM₁; and case VIII – three models invalid.

For the comparison to deep learning with other data fusion strategies, the RF in the present DRFF method was replaced by the support vector machine (SVM) and the K-nearest neighbor (KNN), respectively. As shown in Table 2, different data fusion strategies can achieve similar classification rates. Nevertheless, the presented DRFF exhibits the best classification rate. This means that the RF is a good choice for data fusion of the deep learning features.

The present DRFF method was then compared to deep learning without data fusion. In this category of the comparison, only the AE signal or the vibratory signal were used for the DRFF, or the DBM, respectively. The comparison results are displayed in Table 2. Without data fusion, according to the table, classification rates are dropped. Letting DBM₁ and DBM₂ denote AE signal for DBM, vibratory signal for DBM, respectively, Fig. 5 shows the enhancement of data fusion. As shown in Fig. 5, DRFF, DBM₁ and DBM₂ are all valid (Case I) for 70.88%, only the DRFF and DBM₁ are valid (Case II) for 8.48%, only the DRFF and DBM₂ are valid (Case III) for 16.72%, only the DRFF is valid (Case IV) for 1.82%, only the DBM₁ and DBM₂ are valid (Case V) for 0.05%, only the DBM₂ is valid (Case VI) for 1.09%, only the DBM₁ is valid (Case VII) for 0.35%, and none of them is valid (Case VIII) for 0.61% of the fault diagnosis experiments. This reveals that data fusion enhances the deep learning's performance for the gearbox fault diagnosis.

Moreover, the DRFF method was compared with some “shallow” learning algorithms. The results are also shown in Table 2. The SVM, KNN and RF were directly applied for the fault diagnosis using the AE or the vibratory signals, respectively. In this category, the classification rates are not good. The lowest classification rate occurs at the KNN with the AE signal. Comparing with the KNN, the SVM achieves better classification performance. On the other hand, the RF as a classifier is better than the SVM and the KNN. Compared to deep learning methods (with or without the data fusion), shallow learning methods exhibit inferior performance for the gearbox fault diagnosis. On the other hand, the decrease of the classification rate using the AE signal is higher than when the vibratory signal is used. This means that the vibration signal is more sensitive in the given experiments.

Based on the comparison with different categories of the peer methods, one should have the following points of view: (1) deep learning methods exhibit better performance than shallow learning ones; (2) the AE and the vibratory signals have similar sensitivities for the fault features, while the vibratory sensor exhibits a little bit better than the AE sensor in the experiments; (3) deep learning with a RF classifier is better than the typical deep learning with a simple voting classification

mechanism; (4) as a data fusion tool, the RF is better than the SVM and the KNN; and (5) among all the peer methods, the proposed DRFF technique that utilizes superiorities of both deep learning and the RF shows the best performance for the gearbox fault diagnosis experiments.

It should be noted that the DBM at its current form cannot directly host the vibratory and AE signals. This is different with the image processing field. When we applied the DBM to the sensor signals without the help of the WPT, the extracted features were failed for the fault diagnosis. This means that the standard DBM cannot replace the WPT for the feature extraction. In our further work, we will try to improve the DBM itself to directly extract features of sensor signals instead of WPT statistical parameters.

5. Conclusions

In this paper, a deep random forest fusion (DRFF) method has been reported for the gearbox fault diagnosis using acoustic and vibratory signals, simultaneously. Although the AE sensor and the vibratory accelerometer are all capable of detecting the gearbox defects, their shallow representations are insensitive to distinguish different patterns. In addition, sometimes different diagnosis results are generated by the two different signals. To tackle these problems, the present DRFF exploited two independent DBMs to represent the WPT statistical parameters of the two measurements. An RF was then applied to replace the direct voting classification mechanisms of the two DBMs. In this way, the outputs of the two deep representations were fused together. To evaluate the proposed DRFF method, gearbox fault diagnosis experiments were carried out. The results show that the present method is capable of improving the performance of the gearbox fault diagnosis, comparing with the peer methods.

Acknowledgments

This work is supported in part by the Prometeo Project of the Secretariat for Higher Education, Science, Technology and Innovation of the Republic of Ecuador, the National Natural Science Foundation of China (51375517), and the Youth Science and Technology Talents Project of Chongqing (cstc2014kjc-qncr00003). The valuable comments and suggestions from the two anonymous reviewers are very much appreciated.

References

- [1] D. Galar, A. Thaduri, M. Catelani, L. Ciani, Context awareness for maintenance decision making: a diagnosis and prognosis approach, *Measurement* 67 (2015) 137–150.
- [2] Y. Lei, N. Li, J. Lin, Two new features for condition monitoring and fault diagnosis of planetary gearboxes, *J. Vib. Control* 21 (2015) 755–764.
- [3] J.J. Hou, W.K. Jiang, W.B. Lu, Application of a near-field acoustic holography-based diagnosis technique in gearbox fault diagnosis, *J. Vib. Control* 19 (2013) 3–13.
- [4] A.M.D. Younus, B.S. Yang, Intelligent fault diagnosis of rotating machinery using infrared thermal image, *Expert Syst. Appl.* 39 (2012) 2082–2091.
- [5] J.R. Ottewill, M. Orkisz, Condition monitoring of gearboxes using synchronously averaged electric motor signals, *Mech. Syst. Signal Process.* 38 (2013) 482–498.
- [6] C. Li, M. Liang, Extraction of oil debris signature using integral enhanced empirical mode decomposition and correlated reconstruction, *Meas. Sci. Technol.* 22 (2011) 085701.
- [7] J. Cheng, K. Zhang, Y. Yang, An order tracking technique for the gear fault diagnosis using local mean decomposition method, *Mech. Mach. Theory* 55 (2012) 67–76.
- [8] D. Wang, Q. Miao, R. Kang, Robust health evaluation of gearbox subject to tooth failure with wavelet decomposition, *J. Sound Vib.* 324 (2009) 1141–1157.
- [9] Y. Lei, D. Kong, J. Lin, M.J. Zuo, Fault detection of planetary gearboxes using new diagnostic parameters, *Meas. Sci. Technol.* 23 (2012) 055605.
- [10] B. Li, P.L. Zhang, Q. Mao, S.S. Mi, P.Y. Liu, Gear fault detection using adaptive morphological gradient lifting wavelet, *J. Vib. Control* 19 (2013) 1646–1657.
- [11] J.L.F. Chacon, E.A. Andicoberry, V. Kappatos, G. Asfis, T.H. Gan, W. Balachandran, Shaft angular misalignment detection using acoustic emission, *Applied Acoustics*, 85, pp. 12–22.
- [12] M. Hamel, A. Addali, D. Mba, Investigation of the influence of oil film thickness on helical gear defect detection using Acoustic Emission, *Appl. Acoust.* 79 (2014) 42–46.
- [13] M.R.A.A. Abad, H. Ahmadi, A. Moosavian, M. Khazaei, M.R. Kohan, M. Mohammadi, Discrete wavelet transform and artificial neural network for gearbox fault detection based on acoustic signals, *J. Vibroeng.* 15 (2013) 459–463.
- [14] S. Rafael, A. Romero, S. Souza, T.H. Gan, J. Neasham, G. Goodfellow, Detection of gearbox failures by combined acoustic emission and vibration sensing in rotating machinery, *Insight* 56 (2014) 422–425.
- [15] R.Y. Li, D. He, E. Bechhoefer, Investigation on fault detection for split torque gearbox using acoustic emission and vibration signals, in: *Annual Conference of the Prognostics and Health Management Society*, San Diego, CA, USA, 2009, pp. 1–11.
- [16] S. Souza, P.V. Lieshout, A. Perera, T.H. Gan, B. Bridge, Determination of the combined vibrational and acoustic emission signature of a wind turbine gearbox and generator shaft in service as a pre-requisite for effective condition monitoring, *Renew. Energy* 51 (2013) 175–181.
- [17] M. Khazaei, H. Ahmadi, M. Omid, A. Moosavian, M. Khazaei, Classifier fusion of vibration and acoustic signals for fault diagnosis and classification of planetary gears based on Dempster–Shafer evidence theory, *Proc. Inst. Mech. Eng. E – J. Process Mech. Eng.* 228 (2014) 21–32.
- [18] C. Li, M. Liang, Continuous-scale mathematical morphology-based optimal scale band demodulation of impulsive feature for bearing defect diagnosis, *J. Sound Vib.* 331 (2012) 5864–5879.
- [19] A. Raad, J. Antoni, M. Sidahmed, Indicators of cyclostationarity: theory and application to gear fault monitoring, *Mech. Syst. Signal Process.* 22 (2008) 574–587.
- [20] C. Li, M. Liang, T. Wang, Criterion fusion for spectral segmentation and its application to optimal demodulation of bearing vibration signals, *Mech. Syst. Signal Process.* 64–65 (2015) 132–148.

- [21] F. Chen, B. Tang, R. Chen, A novel fault diagnosis model for gearbox based on wavelet support vector machine with immune genetic algorithm, *Measurement* 46 (2013) 220–232.
- [22] C. Li, M. Liang, Time-frequency signal analysis for gearbox fault diagnosis using a generalized synchrosqueezing transform, *Mech. Syst. Signal Process.* 26 (1) (2012) 205–217.
- [23] C. Li, V. Sanchez, G. Zurita, C.M. Lozada, D. Cabrera, Rolling element bearing defect detection using the generalized synchrosqueezing transform guided by time-frequency ridge enhancement, *ISA Trans.* 60 (2016) 274–284.
- [24] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Science* 313 (2006) 504–507.
- [25] K. Katayama, M. Ando, T. Horiguchi, Model of MT and MST areas using an autoencoder, *Phys. A: Stat. Mech. Appl.* 322 (2013) 531–545.
- [26] T.N. Sainath, B. Kingsbury, G. Saon, H. Soltan, A. Mohamed, G. Dahl, B. Ramabhadran, Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation, *NeuroImage* 108 (2015) 214–224.
- [27] Y. Bai, Z. Chen, J. Xie, C. Li, Daily reservoir inflow forecasting using multiscale deep feature learning with hybrid models, *J. Hydrol.* 532 (2016) 193–206.
- [28] V.T. Tran, F.A. Thobiani, A. Ball, An approach to fault diagnosis of reciprocating compressor valves using Teager–Kaiser energy operator and deep belief networks, *Expert Syst. Appl.* 41 (2014) 4113–4122.
- [29] Z. Chen, C. Li, R.V. Sanchez, Multi-layer neural network with deep belief network for gearbox fault diagnosis, *J. Vibroeng.* 17 (5) (2015) 2379–2392.
- [30] P. Tamilselvan, P. Wang, Failure diagnosis using deep belief learning based health state classification, *Reliab. Eng. Syst. Saf.* 115 (2013) 124–135.
- [31] R. Yan, R.X. Gao, X. Chen, Wavelets for fault diagnosis of rotary machines: a review with applications, *Signal Process.* 96 (2014) 1–15.
- [32] C. Li, R.V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, R.E. Vásquez, Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis, *Neurocomputing* 168 (2015) 119–127.
- [33] J. Schmidhuber, Deep learning in neural networks: an overview, *Neural Netw.* 61 (2015) 85–117.
- [34] L. Shao, D. Wu, X. Li, Learning deep and wide: a spectral method for learning deep networks, *IEEE Trans. Neural Netw. Learn. Syst.* 25 (2014) 2303–2308.
- [35] D.J. Bordoloi, R. Tiwari, Optimum multi-fault classification of gears with integration of evolutionary and SVM algorithms, *Mech. Mach. Theory* 73 (2014) 49–60.
- [36] K. Cho, A. Ilin, T. Raiko, Improved learning of Gaussian–Bernoulli restricted Boltzmann machines, *Lect. Notes Comput. Sci.* 6791 (2011) 10–17.
- [37] Y. Tang, Deep learning using linear support vector machines, in: *Workshop on Representational Learning, ICML 2013, Atlanta, USA, 2013*.
- [38] L. Breiman, Random forests, *Mach. Learn.* 45 (2001) 5–32.
- [39] R. Salakhutdinov, G. Hinton, An efficient learning procedure for deep Boltzmann machines, *Neural Comput.* 24 (2012) 1967–2006.