JOURNAL OF MECHANICAL ENGINEERING AND SCIENCES (JMES)
ISSN: 2289-4659 e-ISSN: 2231-8380
VOL. 16, ISSUE 1, 8653 – 8661
DOI: https://doi.org/10.15282/jmes.16.1.2022.01.0684

ORIGINAL ARTICLE

Gear fault monitoring based on unsupervised feature dimensional reduction and optimized LSSVM-BSOA machine learning model

V.H. Nguyen[1*], V.T. Pham[1]

[1] Faculty of Mechanical Engineering, Hanoi University of Industry, 100000 Hanoi, Vietnam

ABSTRACT – In the trend of Industry 4.0 development, the big data of system operation is significant for analyzing, predicting, or identifying any possible problem. This study proposes a new diagnosis technique for identifying the vibration signal, which combines the feature dimensional reduction method and optimized classifier. Firstly, an auto-encoder feature dimensional reduction (AE-FDR) method is constructed with the bottleneck hidden layer to extract the low-dimensional feature. Secondly, a supervised classifier is formed to carry out fine-turning and classification. The least square-support vector machine (LSSVM) classifier is used as basic with an optimized parameter exploited by the backtracking search optimisation algorithm (BSOA). This LSSVM-BSOA is used to identify the gear fault based on the original vibration data. The proposed AE-FDR-LSSVM-BSOA diagnosis technique shows good ability for identifying the gear fault. A helical gear is experimented with three fault status for evaluate this method. The diagnosis result achieves a high accuracy of 93.3%.

ARTICLE HISTORY
Received: 28th Apr. 2020
Revised: 04th Dec. 2021
Accepted: 06th Dec. 2021

KEYWORDS
Gear fault monitoring;
Feature dimensional reduction;
Auto-encoder AE;
LSSVM-BSOA classification model

INTRODUCTION

In the rotating machinery system, gear status are required for monitoring and identifying its operations for maintenance. This implementation can avoid fatal system breakdowns and prevent economic losses and human safety concerns [1]. Therefore, a reliable diagnosis technique is important in industry applications for monitoring gear health conditions. Normally, the fault diagnosis technique includes the three main phases of data acquisition, feature extraction and pattern classification.

The dimensionality reduction can extract features from the complex vibration data with the goal of improving the generalization performance of the classification phase. In particular, dimensional feature data can cause a computational burden, affect the efficiency of the classification phase due to its time-consumption, and reduce the diagnosis accuracy. The feature dimensional reduction of the original vibration signal can provide useful information for assessing the gear fault condition. The Principal Component Analysis (PCA) [2-4] and Linear Discriminate Analysis (LDA) [5-7] can be effectively used for linear data but not for complex nonlinear and nonstationary vibration data. Recently, deep learning algorithm [8] has attracted research attention for not only processing large scale and high-dimensional datasets [1, 9], but can also learn the feature representation of the nonlinear, nonstationary data [10, 11]. Therefore, in this study, we focus on an auto-encoder (AE) feature dimensional reduction (AE-FDR) method to use the feature dimensionality of the complex vibration data. A deep learning network-based AE architecture with the wise-trained layers is constructed to extract the low-dimensional feature set of the data. Crucially, the unsupervised self-learning algorithm is applied to use a bias/variance tradeoff to mitigate overfitting found in the complex vibration data of gear fault. Regularizing, the low-dimensional obtained feature is a superior way to save space, reduce processing time, and improve diagnostic performance.

Traditionally, classification methods are used to accomplish the diagnostic process in the last phase. These methods have been successfully applied to vibration data such as k-Nearest-Neighbour (k-NN) [12], Artificial Neural Network (ANN) [13-15] and Support Vector Machine (SVM) [4, 13, 16]. Especially, as an improvement of the SVM method, least squares support vector machine (LSSVM) is proposed by Suykens in 1999 [17]. LSSVM can be applied not only to the classification problem but also to cases of regression. LSSVM has been successfully applied in different fields of fault classification. For example, LSSVM is used to identify the fault statuses in gears [18, 19], faults of centrifugal pump [20] and multi-fault diagnosis for rotating machinery [21, 22]. In these case, LSSVM method solution transforms the quadratic programming into linear equations and selects the least square linear system as the loss function. The nonlinear kernel function is used for mapping the low-dimensional inseparable data into high-dimensional separable data. However, the parameters selection of kernel function directly affects the final classification result of LSSVM for solving the nonlinear problem. In this study, we use the backtracking search optimization algorithm (BSOA) to select optimally the parameters of LSSVM, named LSSVM-BSOA. BSOA is a meta-heuristic optimization algorithm and proposed by Civicioglu in 2013 [23]. The BSOA is more efficient, and can solve complex problems [24, 25], especially in the fields of science and engineering [26, 27], and forecasting/ diagnosis [28, 29]. A gained LSSVM-BSOA classifier model takes advantage of regression analysis and generalization performance of the basic LSSVM classifier with the parameters optimized by BSOA. The LSSVM-BSOA model is then used to identify the vibration features of the target data. Finally, a proposed

*CORRESPONDING AUTHOR | V. H. Nguyen | hung2009haui@gmail.com

AE-FDR-LSSVM-BSOA method is formed in combination of AE-FDR and LSSVM-BSOA classifier model to identify the gear fault signal with high accuracy.

The rest of this paper is organized as follows: the AE-FDR method for reducing the feature dimension is proposed. The optimized LSSVM-BSOA classifier model is presented in the next section, which is used to classify the feature pattern of faulty vibration signal. The proposed AE-FDR-LSSVM-BSOA diagnosis technique is applied to gear fault data and the obtained experimental results are analysed and discussed. The conclusion is presented in the last section.

AUTOENCODER FEATURE DIMENSIONAL REDUCTION

AE is a particular architecture of neural network which works as an unsupervised learning algorithm. Figure 1 shows a single AE, which consists of three layers, namely, input, hidden, and output. AE is divided into two stages of the encoder and decoder with the former accomplishing the feature representation from high-dimensional input $x = \{x_1, x_2, \dots, x_n\}$ to low-dimensional data in the hidden layer, $g = \{g_1, g_2, \dots, g_m\}(m \ll n)$ while the latter reconstructs the input x. The input data h back maps the output data $\breve{x} = \{\breve{x}_1, \breve{x}_2, \dots, \breve{x}_n\}$ with high-level feature representation. In Figure 1, layer h is the hidden layer, whereby the inputs are compressed into a small number of neurons. Activation of unit i in layer k is given by Eq. (1).

$$h_i^{(k)} = f\left(\sum_{j=1}^{n} W_{ij}^{(l-1)} x_j^{(l-1)} + b_i^{(1)}\right) \tag{1}$$

where W and b denote weight and bias parameters, respectively. In the first layer, that is, the input layer, $x^{(1)} = x$, and in the last layer, that is, the output layer, $x^{(3)} = \breve{x}$. For the activated function f, we use sigmoid function in hidden layers, but in the output layer, we use linear function because we do not have pre-scale every input example to a specific interval such as $[-1;\ 1]$.

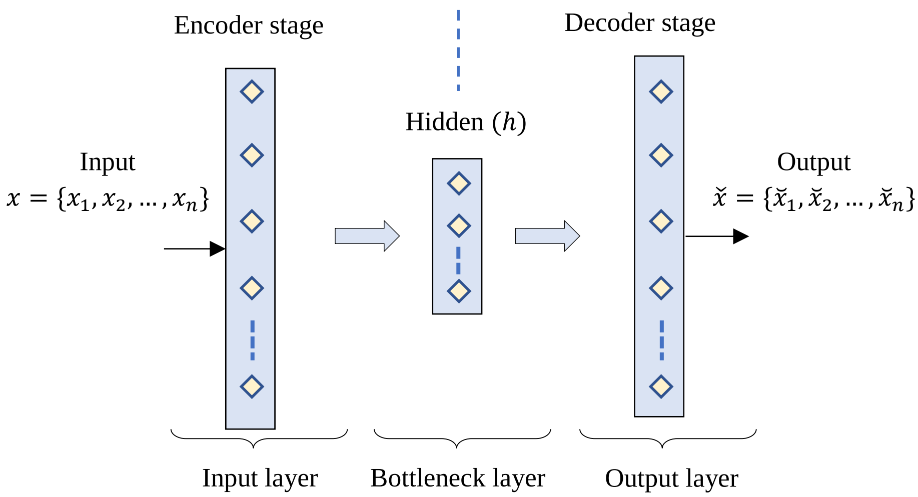

Figure 1. Singular autoencoding architecture

During the training the reconstruction error at the output [8] is minimised using Eq. (2) that shows the objective function with respect to W and b. The objective function includes the regularization term, and the parameter λ determines the strength of regularization.

$$J(W, b) = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{1}{2}\|x(i) - \breve{x}(i)\|^2\right) + \frac{\lambda}{2}\sum_{k=1}^{n_k-1}\sum_{i=1}^{s_k}\sum_{j=1}^{s_k+1}\left(W_{ij}^{(k)}\right)^2 \tag{2}$$

where n_k denotes number of layers in the network and s_k denotes the number of units in the input layer. The idea is to learn an over-complete set of basis vectors to represent input vectors, such that our basis vectors can capture structures and patterns inherent in the input data better. At the same time, to avoid highly compressed encoding that is usually highly entangled, we can encode the input with a small subset of neurons. AE architecture is optimally trained as the data dimensional reduction in hidden layer g. The hidden layer nodes contain most of the important information of the input data that represent as the low- dimensional features.

OPTIMIZED LSSVM-BSOA CLASSIFICATION MODEL

This section presents the important stage of diagnosis. The LSSVM classification method is used as basic classifier with the parameters optimized by the BSOA, named LSSVM-BSOA and the LSSVM-BSOA uses the AE-FDR feature data to train and evaluate.

LSSVM classifier

LSSVM introduces least squares linear system into SVM, which is put forward by Suykens in 1999[17]. LSSVM solves the following convex optimization problems to find the optimal separating hyperplane. Suppose$\{(x_i, y_i)|i = 1,2, \dots, l\}$ is a training set of l sample numbers that corresponds to the category of $y_i \in (-1,1)$, the objective function and constraint condition are shown as follows:

$$\begin{cases} min(J_{LS}) \ with \ J_{LS}(w,e) = \frac{1}{2}w^T w + \frac{1}{2}\gamma \sum_{i=1}^{l} e_i^2 \\ s.t. \ y_i = w^T \varphi(x_i) + b + e_i, \ \ i = 1,2, \dots, l \end{cases} \tag{3}$$

where e_i are slack variables and $\gamma \geq 0$ is a penalty factor/ regularization parameter.

The values of γ influence the training result of the LSSVM model. The low γ value indicates a model with high training errors and the high γ value does not permit any slack variables and consequently increases model complexity. Therefore, finding the proper value for γ is critical and one of LSSVM tuning parameter that should be adjusted conscientiously.

Define Lagrange function:

$$L(w,b,e,\alpha) = \frac{1}{2}w^T w + \frac{1}{2}\gamma \sum_{i=1}^{l} e_i^2 - \sum_{i=1}^{l} \alpha_i (w^T \varphi(x_i) + b + e_i - y_i) \tag{4}$$

where $\alpha_i$is Lagrange multiplier, which can be positive or negative.

The classifying function of LSSVM classifier finally can be obtained as follows:

$$f(x) = sgn\left(\sum_{i=1}^{l} \alpha_i y_i K(x,x_i) + b\right) \tag{5}$$

To obtain better performance and generalization for LSSVM, we use the Gauss radial basic function (RBF) $K(x,y) = exp(-\|x-y\|^2/2\sigma^2)$ to map the samples into a higher dimension feature space. After the LSSVM structure is determined, the parameter pair (γ, σ) that affects the learning performance requires appropriate selection.

Backtracking Search Optimization Algorithm

Backtracking search optimization algorithm (BSOA) is a new meta-heuristic algorithm in the evolutionary series [23]. Information obtained from past generations is used to search for better fitness solutions. Table 1 presents the general structure of BSOA. The bio-inspired philosophy of BSOA corresponds with a social group of living creatures, which relates to hunting areas for finding food at random movements.

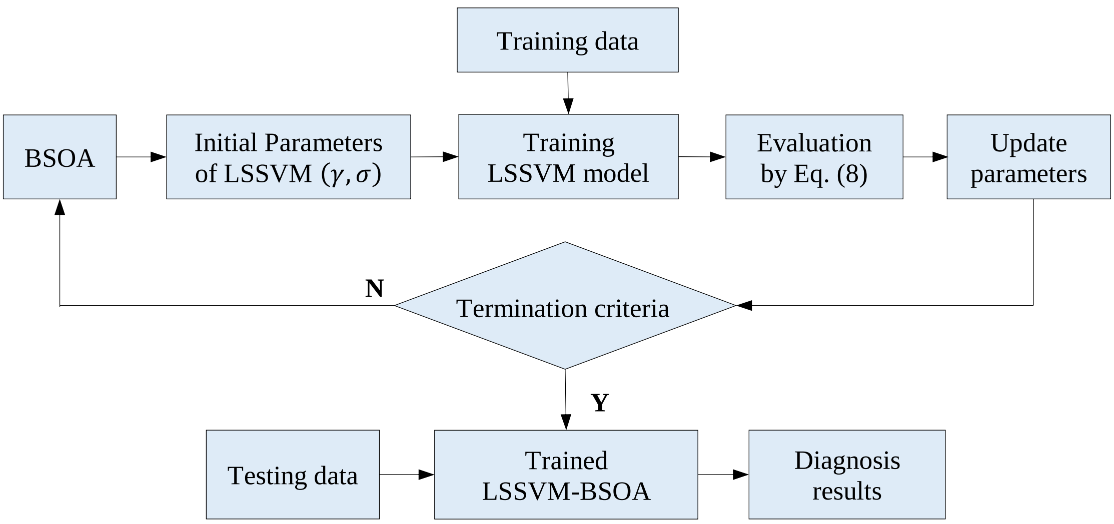

Figure 2. Constructing the optimized LSSVM-BSOA classification model

Table 1. Backtracking search optimization algorithm (BSOA)

Initialization
Repeat
Selection I
Generate new population:
Mutation
Crossover
End
Selection II
Until the stop conditions are met

In the initialization phase, the algorithm generates and evaluates the initial population P and starts a historical population P_o . The historical population constitutes the memory of BSOA.

In the Selection I, the algorithm randomly determines whether the current population P is recorded as the historical population P_o. The individuals of P_o are then shuffled.

The mutation operator creates P_m , which forms an initial version of the new population P_n, according to Eq. (6). Therefore, P_m is the result of the movement of individuals of P in the directions set by $(P_o - P)$. Then, F defines the motion amplitude, and is given by Eq. (7).

$$P_m = P_n + F(P_o - P) \tag{6}$$

$$F = k \cdot r_{rand} \tag{7}$$

where k value is adjusted empirically during prior simulations; $r_{rand} \sim N(0,1)$, N is the standard normal distribution. The operator randomly crosses over elements from P_m with elements of P, which generates the final version of P_n

In the Selection-II, the algorithm selects elements from P_n to have better fitness than elements of P. Then P_n replaces them in P. Thus, P only receives newly evolved individuals. After meeting the stopping conditions, the algorithm returns the best solution found.

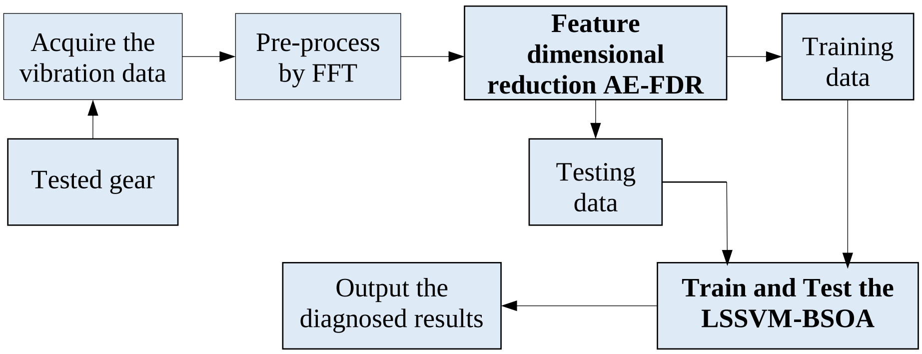

Figure 3. Flowchart of the proposed LSSVM-BSOA classifier

Optimised LSSVM-BSOA Method

The LSSVM parameters (γ, σ) play a key role in constructing the model that can be obtained using an optimization algorithm. BSOA is used to explore the search space of the given LSSVM classification problem to find the parameters (γ, σ) required to maximize the particular objective of accuracy [11]. The principled training phase of the optimal LSSVM-BSOA classifier model includes several main steps, which is implemented as follows:

Step 1: Prepare the training feature data using the AE-FDR method.

Step 2: Initiate LSSVM with the random parameters γ and σ. Set iterative variable $t = 0$ and perform the training process for the next steps.

Step 3: Increasing iteration variable by set $t = t + 1$.

Step 4: Deterioration evaluation. The fitness function f is used to evaluate the quality of every element, related to the Eq. (8)

$$f(\%) = \frac{y_{classified}}{N} 100 \tag{8}$$

where $y_{classified}$ is the right classified samples, and N is the summed samples in the testing data. The desirable absolute value is large and classification accuracy is high.

Step 5: Stop criteria checking. The deterioration function satisfies Eq. (8) or iteration is maximal go to step 7. If not, go to the next step.

Step 6: Update the new γ and σ parameters, go to step 3.

Step 7: End of training, the trained LSSVM-BSOA classifier model is optimized.

The efficient search capability of BSOA incorporates the generalization capability of LSSVM that can bring out the higher classification accuracy. Figure 2 shows the architecture for LSSVM-BSOA. Each reactant represents the candidate solution for the model that includes penalty parameter γ and the kernel function parameter σ.

PROPOSED FAULT DIAGNOSIS TECHNIQUE FOR GEARS

Proposed Technique with AE-FDR- LSSVM-BSOA

The AE-FDR- LSSVM-BSOA proposed technique is formed including two main stages: (1) AE-FDR dimensional reduction method extracts the feature of gear fault vibration data and (2) LSSVM-BSOA is the classification model. Firstly, AE-FDR exploits the features from the original vibration signal that means increasing the overall reliability and identification accuracy of the classification model. Secondly, a classifier LSSVM-BSOA model is optimized to accurately identify the fault statuses of health gear (HG), missed tooth (MT), and broken tooth (BT), respectively. This technique can be visualized by the flowchart in Figure 3, which can be described as follows:

Step 1: Collect the vibration data of the gear box with tested opinion gear

Step 2: Use the FFT to pre-processing the data

Step 3: Reduce the feature dimension using AE-FDR method. Then, divide the feature data into training-testing parts

Step 4: Train and test the AE-FDR- LSSVM-BSOA technique to identify the actual gear fault. The gear fault diagnostic technique based on AE-FDR- LSSVM-BSOA demonstrate diagnostic accuracy and reliability in experiments.

Experimental Analysis

To evaluate the proposed diagnosis technique, we use a test rig shown in Figure 4 to generate the experimental vibration data. Three gear fault conditions on driving shaft are considered HG class, MT class, and BT class as shown in Figure 5. The main gear parameters are teeth number16, pressure angle 8^{o} and module 2.5. An acceleration sensor of Triaxial Delta Tron 4525B-001 type (Bruel&Kjær) is fixed on the bearing house with sample frequency of 1024 Hz and rotating frequency of 35Hz. A group of 45 vibration signals is collected following three classes of gear fault status with 15 signals from each tested gear condition, respectively. Table 2 describes the acquired vibration data, in which 30 samples are used for training the classification model and the rest of data is used to evaluate the diagnosis technique. Figure 6 shows a vibration signal sample in the time domain with three tested gear faults, which may imply that the vibration intensity is unclear and cannot be determined even though the loading parameter is unchanged.

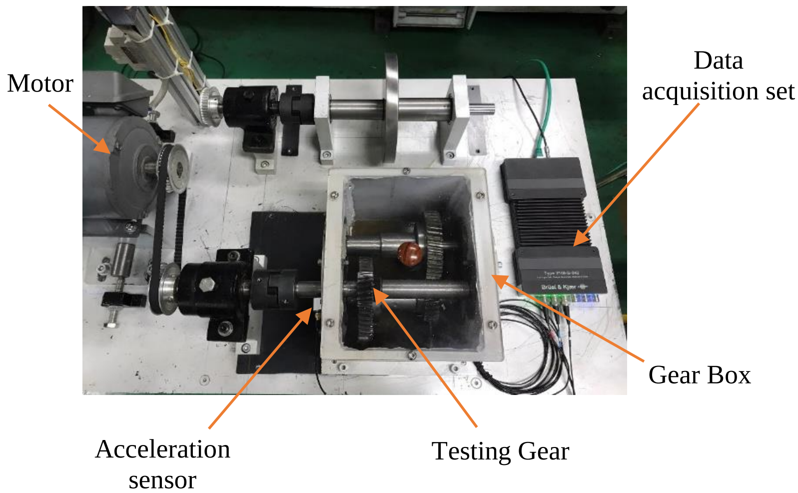

Figure 4. Schematic of the experimental setup

Figure 5. Testing gear conditions: a) Health gear (HG); b) Missed tooth (MT) andc) Broken tooth (BT)

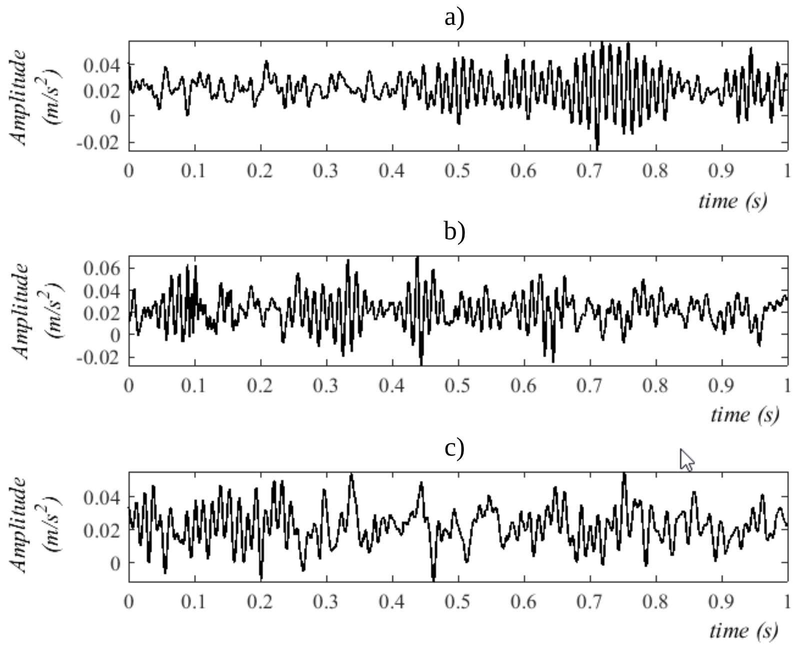

Figure 6. Sample of gear vibration data at statuses of (a) HG state; (b) MT state and (c) BT state

Firstly, implementing, the obtained vibration data is pre-processing using the FFT method to present the frequency domain characteristics. Secondly, the AE-FDR is constructed with layer-wise training of bottleneck hidden layer to reduce the feature dimension as a fault feature extraction from the vibration data. For the auto-encoder, we constructed the network architecture of $1024 - d - 1024$ (d is the reduced feature dimension) and adjust the parameters in the objective function Eq. (3) as $\lambda = 0.001, \beta = 6$. Thirdly, the sensitive fault feature matrix is used to train the optimal LSSVM-BSOA classification model. Lastly, the sensitive fault features including five values of each testing data are extracted. The 15 sensitive values of fault feature are input to the trained LSSVM-BSOA model with the outputs of testing data. Table 3 shows the results of training time and the identification rates.

Table 2. Collection of vibration signal samples

Gear conditions	Class	Training data (*sample*)	Testing data (*sample*)
Health Gear (HG)	1	15	5
Missed Tooth (MT)	2	15	5
Broken Tooth (BT)	3	15	5

Table 3. AE-FDR-LSSVM-BSOA method diagnosis results

Class	Testing data	d = 3	d = 4	d = 5	d = 6	d = 7	d = 8
1	$(x_{16} - x_{20})$	1(3)3(2)	1(5)	1(5)	1(3)3(2)	1(4)3(1)	1(3)3(2)
2	$(x_{36} - x_{40})$	2(4)3(1)*	2(1)1(1)3(1)	2(4)3(1)	2(3)1(1)3(1)	2(4)3(1)	2(4)3(1)
3	$(x_{56} - x_{60})$	3(5)	3(5)	3(5)	3(5)	3(4)	3(5)
Diagnosis rate (%)		80.0	86.7	93.3	73.3	86.7	80.0
Training time (s)		0.92	1.16	1.49	1.93	2.89	4.62

** Represents that among the 5 testing sample, 4 are diagnosed as the 2th class and 1 is misclassified to 3rd class, others are similar to this*

For comparison, the feature dimension ($d = 3, 4, \ldots, 8$) reduced by AE-FDR is constructed as the sensitive fault feature vector. The obtained features are input to the LSSVM-BSOA classifier model with the same number of training and testing samples. Figure 7 presents the scatter of the feature dimension d_i. The implementation is the same as the output results of testing data, and the identification rates are given in Table 3. The results show that the proposed method with $d = 5$ has a good classifying result for testing data with a high identification rate (93.3%), which is suitable to scatter in Figure 7 (c). The feature values ($d = 5$) can clearly distinguish the mapping and accurately separate the gear fault status.

The results in Table 3 also show that the input features influent directly to the fault diagnosis results. The fault feature dimensioning five values ($d = 5$) is much more suitable for diagnosing the gear fault. If the dimension of the input fault feature is low ($d = 3,\ 4$), the obtained features cannot entirely reflect the gear fault information as the MT of class 2 of an insensitive status and lastly cause a low efficiency (80% diagnosis rate in $d = 3$). If the feature dimension is high ($d = 6,\ 7,\ and\ 8$), the obtained features can contain the redundant information, and cause low classification accuracy. In addition, the high dimension consumes time for training, as the diagnosis rate 80% in $d = 8$ consumes 4.62 (s). Thus, we recommend that $d = 5$ may be more suitable to obtain the better result of classification accuracy. This is the key of the proposed diagnosis method.

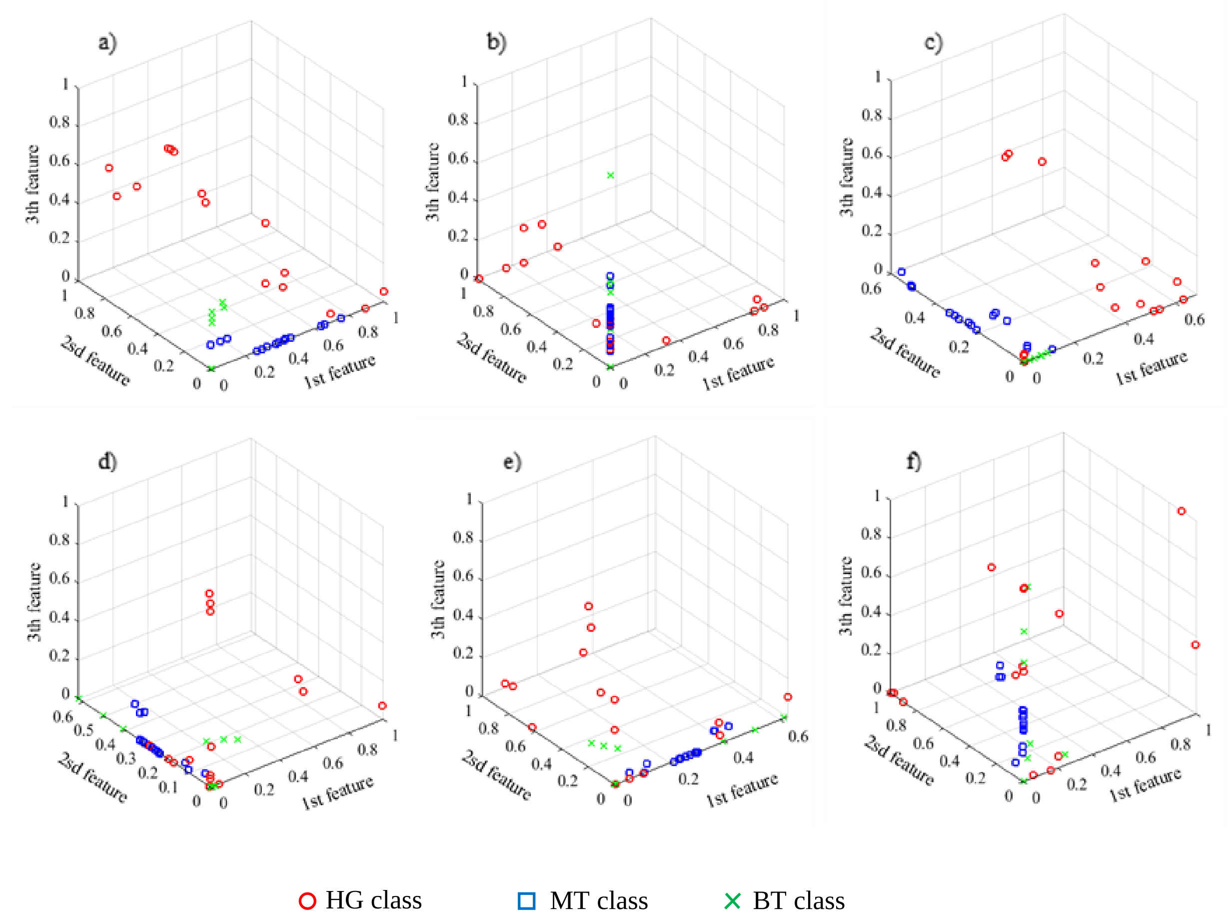

Figure 7. Scatter plot for the reduced feature set, (a) $d = 3$; (b) $d = 4$; (c) $d = 5$; (d) $d = 6$; (e) $d = 7$; (f) $d = 8$

CONCLUSIONS

In this study, we present an expert study at the gear fault identification based on the integrated AE-FDR method and LSSVM-BSOA classifier model. We explore the suitability of deep learning AE-FDR network architecture into the feature dimensional reduction. Based on the complex the vibration data, the low-dimensional crucial feature set with five values ($d = 5$) are extracted at the AE-FDR bottleneck layer. These features are the most important fault information closely related to the gear status, which effectively provides for fault diagnosis tasks. The optimized LSSVM-BSOA classifier model is formed to identify the gear fault status based on the obtained features, which confirms its superiority. The experimental vibration data of the three different gear faults are used to demonstrate that the proposed AE-FDR-LSSVM-BSOA technique has good ability and is highly effective in accurate fault diagnosis. Based on these achievements, we improve the AE-FDR- LSSVM-BSOA ability to recognize the sensitive faults in the rotating machinery part without the training data and apply the method to predict their remaining useful life.

ACKNOWLEDGMENTS

The authors also acknowledge the support of the Faculty of Mechanical Engineering, Hanoi University of Industry, Hanoi, Vietnam.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

[1] Z. Chen, S. Deng, X. Chen, C. Li, R.-V. Sanchez, and H. Qin, "Deep neural networks-based rolling bearing fault diagnosis," *Microelectronics Reliability*, vol. 75, pp. 327-333, 2017, doi:10.1016/j.microrel.2017.03.006.

[2] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 228-233, 2001, doi:10.1109/34.908974.

[3] A. Prieto-Moreno, O. Llanes-Santiago, and E. García-Moreno, "Principal components selection for dimensionality reduction using discriminant information applied to fault diagnosis," *Journal of Process Control*, vol. 33, pp. 14-24, 2015, doi:10.1016/j.jprocont.2015.06.003.

[4] V. H. Nguyen, J. S. Cheng, and V. T. Thai, "An integrated generalized discriminant analysis method and chemical reaction support vector machine model (GDA-CRSVM) for bearing fault diagnosis," *Advances in Production Engineering & Management*, vol. 12, pp. 321-336, 2017, doi:10.14743/apem2017.4.261.

[5] L. Zhang, Q. Zhang, L. Zhang, D. Tao, X. Huang, and B. Du, "Ensemble manifold regularized sparse low-rank approximation for multiview feature embedding," *Pattern Recognition*, vol. 48, pp. 3102-3112, 2015, doi:10.1016/j.patcog.2014.12.016.

[6] B. Yao, P. Zhen, L. Wu, and Y. Guan, "Rolling element bearing fault diagnosis using improved manifold learning," *IEEE Access*, vol. 5, pp. 6027-6035, 2017, doi:10.1109/ACCESS.2017.2693379.

[7] M. Zhao, B. Tang, and Q. Tan, "Bearing remaining useful life estimation based on time–frequency representation and supervised dimensionality reduction," *Measurement*, vol. 86, pp. 41-55, 2016, doi:10.1016/j.measurement.2015.11.047.

[8] Y. Bengio, "Learning Deep Architectures for AI," *Foundations and Trends® in Machine Learning,"* vol. 2, pp. 1-127, 2009, doi:10.1561/2200000006.

[9] P. Hou, C. Wen, and D. Dong, "Rolling bearing fault diagnose based on stacked sparse auto encoder," in *2017 36th Chinese Control Conference (CCC)*, 2017, pp. 7027-7032.

[10] V. H. Nguyen, J. S. Cheng, Y. Yu, and V. T. Thai, "An architecture of deep learning network based on ensemble empirical mode decomposition in precise identification of bearing vibration signal," *Journal of Mechanical Science and Technology*, vol. 33, pp. 41-50, 2019, doi:10.1007/s12206-018-1205-6.

[11] V. Nguyen, T. D. Hoang, V. Thai, and X. Nguyen, "Big vibration data diagnosis of bearing fault base on feature representation of autoencoder and optimal LSSVM-CRO classifier model," in *2019 International Conference on System Science and Engineering (ICSSE)*, 2019, pp. 557-563.

[12] D. Dou and S. Zhou, "Comparison of four direct classification methods for intelligent fault diagnosis of rotating machinery," *Applied Soft Computing*, vol. 46, pp. 459-468, 2016, doi:10.1016/j.asoc.2016.05.015.

[13] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Fault diagnosis of ball bearings using machine learning methods," *Expert Systems with Applications*, vol. 38, pp. 1876-1886, 2011, doi:10.1016/j.eswa.2010.07.119.

[14] J. P. Patel and S. H. Upadhyay, "Comparison between artificial neural network and support vector method for a fault diagnostics in rolling element bearings," *Procedia Engineering*, vol. 144, pp. 390-397, 2016, doi:10.1016/j.proeng.2016.05.148.

[15] Y. Yu, YuDejie, and C. Junsheng, "A roller bearing fault diagnosis method based on EMD energy entropy and ANN," *Journal of Sound and Vibration*, vol. 294, pp. 269-277, 2006, doi:10.1016/j.jsv.2005.11.002.

[16] J. Ben Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, and F. Fnaiech, "Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals," *Applied Acoustics*, vol. 89, pp. 16-27, 2015, doi:10.1016/j.apacoust.2014.08.016.

[17] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, pp. 293-300, June 01 1999, doi:10.1023/a:1018628609742.

[18] X. Liu, L. Bo, and H. Luo, "Bearing faults diagnostics based on hybrid LS-SVM and EMD method," *Measurement*, vol. 59, pp. 145-166, 2015, doi:10.1016/j.measurement.2014.09.037.

[19] Y. Zhang, Y. Qin, Z.-y. Xing, L.-m. Jia, and X.-q. Cheng, "Roller bearing safety region estimation and state identification based on LMD–PCA–LSSVM," *Measurement*, vol. 46, pp. 1315-1324, 2013, doi:10.1016/j.measurement.2012.11.048.

[20] Z. Yunlong and Z. Peng, "Vibration fault diagnosis method of centrifugal pump based on EMD complexity feature and least square support vector machine," *Energy Procedia*, vol. 17, pp. 939-945, 2012, doi:10.1016/j.egypro.2012.02.191.

[21] Z. Su, B. Tang, Z. Liu, and Y. Qin, "Multi-fault diagnosis for rotating machinery based on orthogonal supervised linear local tangent space alignment and least square support vector machine," *Neurocomputing*, vol. 157, pp. 208-222, 2015, doi:10.1016/j.neucom.2015.01.016.

[22] V. Nguyen, T. Hoang, V. Thai, Q. Nguyen, and X. Nguyen, "Identification of gear fault signal based on adaptive EMD feature extraction and optimal GA-LSSVM classification model," *Cham*, 2020, pp. 406-418.

[23] P. Civicioglu, "Backtracking search optimization algorithm for numerical optimization problems," *Applied Mathematics and computation,* vol. 219, pp. 8121-8144, 2013.

[24] B. A. Hassan and T. A. Rashid, "Operational framework for recent advances in backtracking search optimisation algorithm: A systematic review and performance evaluation," *Applied Mathematics and Computation,* p. 124919, 2019.

[25] A. O. de Sá, N. Nedjah, and L. de Macedo Mourelle, "Genetic and backtracking search optimization algorithms applied to localization problems," in *Computational Science and Its Applications – ICCSA 2014*, Cham, 2014, pp. 738-746.

[26] S. Wang, X. Da, M. Li, and T. Han, "Adaptive backtracking search optimization algorithm with pattern search for numerical optimization," *Journal of Systems Engineering and Electronics,* vol. 27, pp. 395-406, 2016.

[27] F. Zou, D. Chen, S. Li, R. Lu, and M. Lin, "Community detection in complex networks: Multi-objective discrete backtracking search optimization algorithm with decomposition," *Applied Soft Computing,* vol. 53, pp. 285-295, 2017.

[28] V. Thai, J. Cheng, V. Nguyen, and P. Daothi, "Optimizing SVM's parameters based on backtracking search optimization algorithm for gear fault diagnosis," *Journal of Vibroengineering,* vol. 21, pp. 66-81, 2019, doi:10.21595/jve.2018.19859.

[29] Z. Wang, Y.-R. Zeng, S. Wang, and L. Wang, "Optimizing echo state network with backtracking search optimization algorithm for time series forecasting," *Engineering Applications of Artificial Intelligence,* vol. 81, pp. 117-132, 2019.