

A Robust Approach of Multi-sensor Fusion for Fault Diagnosis Using Convolution Neural Network

Jiahao Sun,¹ Xiwen Gu,¹ Jun He,¹ Shixi Yang,¹ Yao Tu,² and Chenfang Wu¹

¹State Key Laboratory of Fluid Power and Mechatronics Systems, Zhejiang University, Hangzhou, China

²Hangzhou Steam Turbine & Power Group Co., Ltd, Hangzhou, China

(Received 29 March 2022; Revised 31 May 2022; Accepted 06 June 2022; Published online 07 June 2022)

Abstract: Multi-sensor measurement is widely employed in rotating machinery to ensure the safety of machines. The information provided by the single sensor is not comprehensive. Multi-sensor signals can provide complementary information in characterizing the health condition of machines. This paper proposed a multi-sensor fusion convolution neural network (MF-CNN) model. The proposed model adds a 2-D convolution layer before the classical 1-D CNN to automatically extract complementary features of multi-sensor signals and minimize the loss of information. A series of experiments are carried out on a rolling bearing test rig to verify the model. Vibration and sound signals are fused to achieve higher classification accuracy than typical machine learning model. In addition, the model is further applied to gas turbine abnormal detection, and shows great robustness and generalization.

Key words: deep learning; engineering application; fault diagnosis; multi-sensor fusion

I. INTRODUCTION

Rotating machinery predisposes to various faults under extreme working speed and load for a long time, which directly threaten the safe operation of mechanical systems. Undetected faults may lead to undesirable vibration and noise, the reduction of equipment service life, huge economic losses, or even catastrophic personal casualty. A reliable and robust approach to machine diagnosis will detect the fault in time and reduce the cost of maintenance. For the guarantee of the safe operation of machine equipment with large size and complex structure, a large number of sensors are mounted to monitor whether the equipment works. Multi-sensor data provides more complementary information of the health condition, which helps achieve higher diagnosis accuracy [1].

In order to extract the valid information and reduce the redundant information from a mass of sensors, the approaches of fusion were proposed. Depending on which signal processing step the information is combined together, common fusion methods can be classified as data-level fusion, feature-level fusion, and decision-level fusion [2]. Inturi [3] extracted the statistical features of the vibration and sound signals based on EMD and identified the fault sensitivity features by utilizing the decision tree algorithm. Safizadeh et al. [4] employed principal component analysis (PCA) to reduce the redundancy of features, after calculating the vibration and load signal indices. Gomes et al. [5] used a recursive feature elimination algorithm to screen features in the time and frequency domain of vibration and sound signals, and then established the mapping relationship between the optimal features and the tool states. Tang et al. [6] used the multi-layer selective integration algorithm to simulate the cognitive process of experts and fused the feature information of sound and vibration spectrum by integration sub-models. Dewallef et al. [7] integrated a Kalman filter and a Bayesian belief

networks for fault diagnosis, and achieved decision-level fusion of information. Based on fuzzy measure and fuzzy integral theory, Liu et al. [8] proposed a feature-level fusion model and a decision-level fusion model for fault diagnosis. However, with the deepening of fusion level, the complexity of the method increases. In contrast, data-level fusion is used to combine the information from multi-sensor before the complex feature extraction operation, less prior knowledge is required. Common fusion methods includes weighted average [9–11], machine learning (ML) and so on.

ML-based method has been widely used for multi-sensor fusion in engineering [12]. Guo [13] converted accelerometer signals to several time-frequency feature maps using continuous wavelet transform (CWPT), and the maps were fused by a convolution neural network (CNN) later. Lu et al. [14] employed empirical mode decomposition (EMD) and multi-fractal to fuse signals and construct the feature maps. Some simple geometric operation is also used to construct the input map, without the advanced signal processing. Zhang et al. [15] fused accelerometer signals by splicing their spectrum and using deep boltzmann machine (DBN) to identify ball screw performance degradation. Azamfar et al. [16] stacked the motor current signals into a matrix as input. Wang et al. [17] extracted features of accelerometer signals and built a composite image as input. These methods using signals with the same physical quantity. However, signals with many physical quantities are used in engineering, including vibration, sound, air pressure, flow, etc. The research of a robust and adaptive fusion method is necessary.

This paper proposed a robust MF-CNN method for machine fault diagnosis, which executes the fusion process by a 2-D convolution layer. The principal contributions of this paper are summarized as follow:

- (1) An effective multi-sensor fusion method is proposed, which could execute multi-sensor signals fusion to improve the diagnosis accuracy. Series of tests are conducted to validate the method.

Corresponding author: hjshenhua@zju.edu.cn

- (2) A robust MF-CNN model is established which has great generalization and is robust to different industrial application. Moreover, it performs well for bearing fault diagnosis with strong noise disturbing.
- (3) The proposed fault diagnosis model could execute fault diagnosis without prior knowledge or advanced data pre-processing methods.

Below is an outline of the remainder of the paper. The theoretical background of this approach is introduced in section II, and the MF-CNN approach is introduced in section III. In section IV, the performance of the model is valid by several experiments. And the robustness and generalization of the model is valid by two applications in engineering. Finally, conclusions are drawn in section V.

II. BACKGROUND

A. BASIC COMPONENTS OF CNN

CNN is a typical deep feed-forward network model. The basic CNN frameworks usually include several convolutional layers, pooling layers, activation layers and fully-connected layers [18].

The convolutional layer is the main component of CNN, which aims to learn feature of the inputs. Convolutional layer consists of several convolution filters to extract different kinds of feature. To generate complete feature maps, the weight of kernels will be shared by all spatial locations of the input. The output value $z_{i,j,k}^l$ is calculated by:

$$z_{i,j,k}^l = w_k^l x_{i,j}^l + b_k^l \quad (1)$$

where index (i, j, k) means the position (i, j) of the k -th matrix, l means the layer number. w is the weight vector, b is the bias, and the x is the input matrix. Particularly, convolution operation is a process of weighted addition of values of multi-sensor signals at the same index.

The pooling layer is usually inserted between two convolutional layers. It could reduce the resolution of the feature maps for faster convergence. Denoting the pooling function as $\text{pool}(\cdot)$, which commonly represent max pooling or average pooling function, for local feature map $a_{i,j,k}^l$, we have:

$$y_{i,j,k}^l = \text{pool}(a_{i,j,k}^l) \quad (2)$$

where y is the calculation result of the pooling layer.

ReLU is one of the most famous activation functions, which converts the negative part to zero. The ReLU function could be concluded as:

$$a_{i,j,k} = \max(z_{i,j,k}, 0) \quad (3)$$

where z is the input of the activation function. ReLU allows the network to solve more nonlinear problems, calculate faster than other activation functions, and easily obtain sparse representations.

After stacking several layers for feature extraction, there might be a fully-connected layer for final reason and classification. And the last layer of CNN is output layer. The Softmax function is usually used for classification tasks:

$$S_i = \frac{e^i}{\sum_j e^j} \quad (4)$$

where S_i is the probability of the i -th class among j class.

It is important to choose an appropriate loss function for a specific task. During training, the optimum parameters can be obtained by minimizing the loss of the network. Denoting N desired output $o^{(n)}$ of CNN with target label $y^{(n)}$, where n means the n -th output, the loss could be calculated by:

$$L = \frac{1}{N} \sum_{n=1}^N l(\theta; y^{(n)}, o^{(n)}) \quad (5)$$

where θ represents the optimum parameters of the loss function.

B. OPTIMIZATION

The ability of trained deep CNNs is highly dependent on the amount of training data. Data augmentation [19] is a straightforward strategy to solve data scarcity, which transforms the available data in a simple geometric way without changing their features. For a 1-D time serial, shifting is the operation commonly used. A fixed window is shifted across the data with partial overlap.

The backpropagation algorithm uses gradient descent to update the parameters. Standard gradient descent is conducted by:

$$\theta_{t+1} = \theta_t - \eta \nabla_{\theta} E[L(\theta)] \quad (6)$$

where $E[L(\theta)]$ is the expectation of loss value over the batch of training data, and η is the learning rate. Nesterov momentum [20] is a gradient descent method which takes

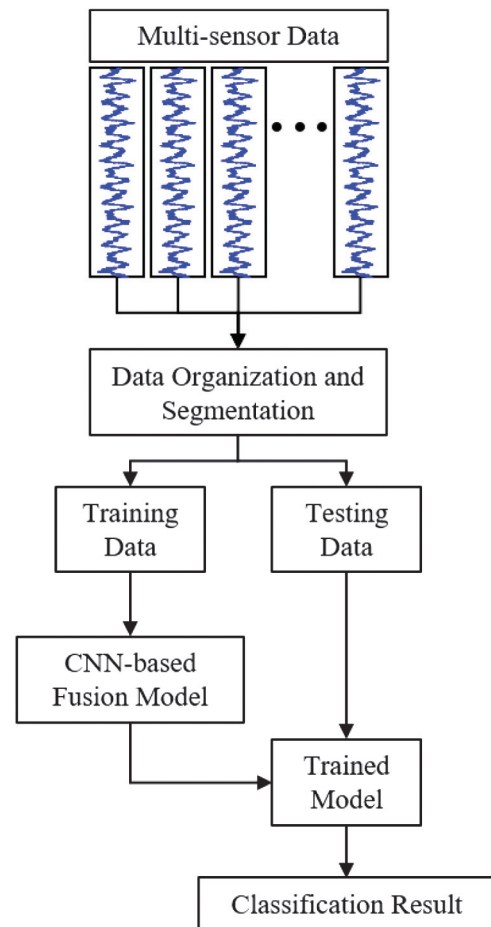


Fig. 1. Flow chart of the multi-sensor fusion method.

historical gradient into consideration and anticipatory update to prevent the excessive optimization. Nesterov momentum can be described as:

$$v_{t+1} = \gamma v_t - \eta_t \nabla_{\theta} L(\theta_t + \gamma v_t; x^{(t)}, y^{(t)}) \quad (7)$$

$$\theta_{t+1} = \theta_t + v_{t+1} \quad (8)$$

where v_{t+1} is the current velocity vector, and γ is the momentum term.

With the increase of network depth, the distribution of feature maps will greatly differ from the raw data, which may result in an inaccurate network. BN layer linearly transforms batch of data to have zero-mean and unit variance to alleviate this phenomenon.

III. PROPOSED MULTI-SENSOR FUSION METHOD BASED ON CNN

Figure 1 shows the flowchart of the proposed multi-sensor fusion method. Signals captured by multi-sensor have redundant and complementary information, which may contain the health information of the machine. It is important to fuse multi-sensor signals with appropriate weights to keep the complementary information and abandon the redundant information.

Multi-sensor signals are a combination of multiple one-dimensional signals, which can be arranged into a two-dimensional array. One dimension represents time, and the other dimension is the number of channels. The first approach that comes to mind should be to directly use 2D-CNN for classification. Theoretically, 1-D CNN is a special form of 2-D CNN. However, multi-sensor signals have physical meaning in the time dimension, and have no physical meaning in the channel number dimension. The direct use of 2-D CNN may attach some redundant meaning to them. The sequence of multi-sensor signals forming the matrix will also influence the diagnosis results. The method adopted in this paper is to use a convolution operation to fuse the signals of all channels into a 1-D signal to avoid this problem.

In this study, MF-CNN method is proposed for machine fault diagnosis. The structure of MF-CNN model is shown in Fig. 2, including a fusion layer, multiple convolutional layers following a pooling layer and a BN layer, and finally a fully-connected layer. Signals in time serial collected by sensors are stacked row by row to construct a 2-D matrix as input, and then the fusion layer fuses the 2-D matrix to a 1-D vector via convolution operation. The weight matrix can be trained with the whole model, and reaches a better value through training. Multiple convolutional layers and pooling layers aims to automatically extract features, and fully-connected layer performs advanced reasoning and final classification.

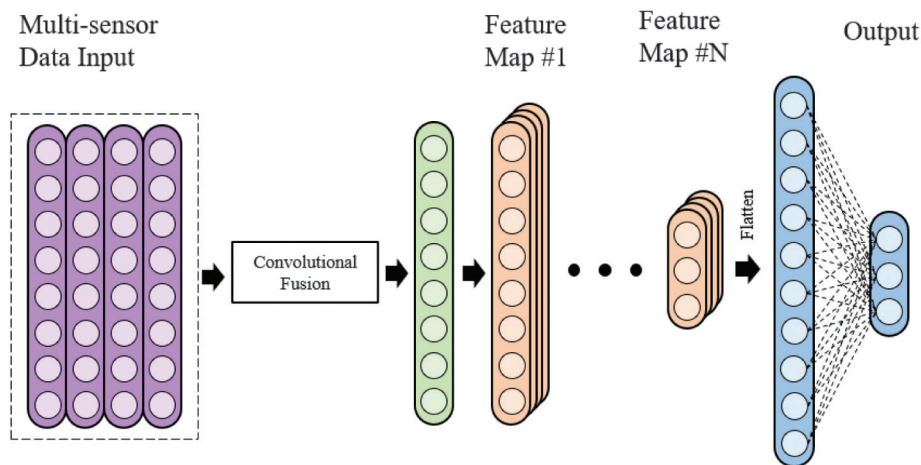


Fig. 2. The proposed fusion CNN architecture.

Table I Parameters of the fusion CNN model

Layer Name	Layer Type	Layer Description
Fusion Layer	Convolution	N = 1, shape = [6, number of channels], stride = 1, padding = 'valid', activation = 'ReLU'
Conv #1	Convolution	N = 16, shape = [1,4], stride = 1, activation = 'ReLU'
Pooling #1	Maxpooling	Pool size = [1,2], stride = 2
Conv #2	Convolution	N = 32, shape = [1,4], stride = 1, activation = 'ReLU'
Pooling #2	Maxpooling	Pool size = [1,2], stride = 2
Conv #3	Convolution	N = 32, shape = [1,6], stride = 1, activation = 'ReLU'
Pooling #3	Maxpooling	Pool size = [1,2], stride = 2
Conv #4	Convolution	N = 32, shape = [1,8], stride = 1, activation = 'ReLU'
Pooling #4	Maxpooling	Pool size = [1,2], stride = 2
Output	Fully Connected	NN = 4 (4 classes output), activation = 'Softmax'

N = number of filters; NN = number of nodes. Maximum epoch is 50 and the learning rate is set to 0.05.

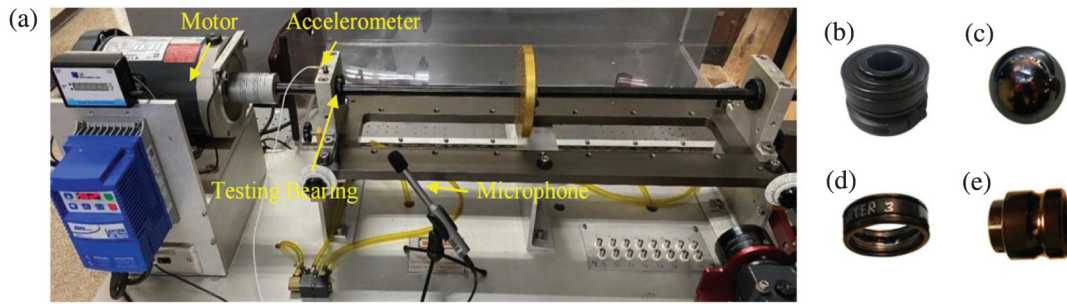


Fig. 3. The experiment equipment (a) rolling bearing test rig. (b) healthy bearing. (c) rolling element fault. (d) outer race fault. (e) inner race fault.

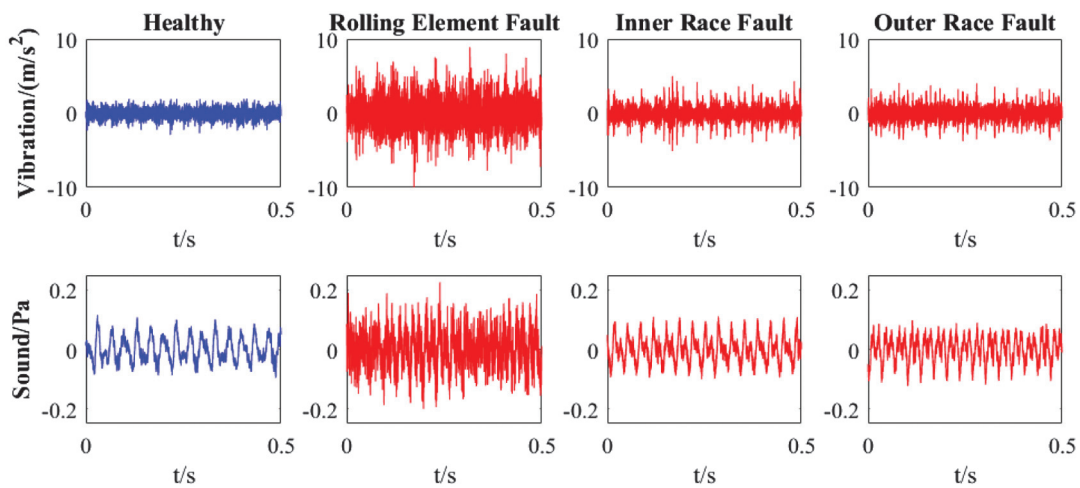


Fig. 4. A comparison between raw signals collected at 1800 rpm from a bearing in four health conditions including Healthy, Rolling Element Fault, Inner Race Fault and Outer Race Fault.

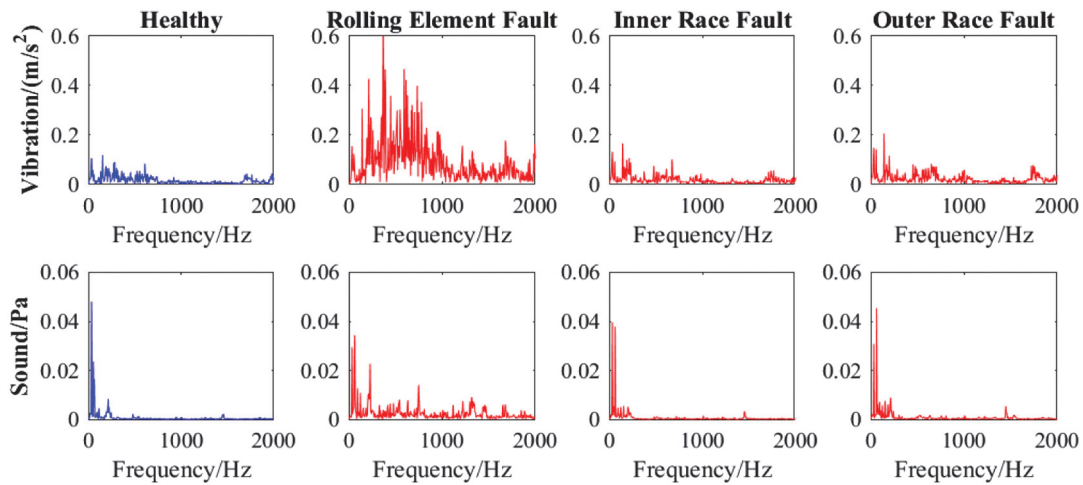


Fig. 5. A comparison between spectra of signals collected at 1800 rpm from bearings in four health conditions, including Healthy, Rolling Element Fault, Inner Race Fault and Outer Race Fault.

Table II The details of the dataset

Bearing condition	Working speeds	Length of samples	Number of samples
Healthy	600, 1200, 1800, 2400 rpm	4800	1000
Rolling Element Fault	600, 1200, 1800, 2400 rpm	4800	1000
Outer Race Fault	600, 1200, 1800, 2400 rpm	4800	1000
Inner Race Fault	600, 1200, 1800, 2400 rpm	4800	1000

To achieve a fast and stable convergence, learning rate schedules and classical momentum optimizers are used. The parameters of the proposed model are shown in Table I. It should be noted that the proposed model is able to fuse multi-sensor signals automatically in data level, so that the loss of information is minimized.

IV. EXPERIMENT VALIDATION

A. EXPERIMENT ON ROLLING BEARING TEST RIG

Rolling elements bearing is the important part of rotating machinery systems, and health conditions of bearings directly affect whether the machine can work properly. In this work, an experiment on rolling bearing fault diagnosis is conducted for method validation. Fig. 3a shows the test rig used in experiment. The rotor is driven by a motor (0.735 kW, three-phase) and a data acquisition system is used to collect the data. The driven speed range of the motor is 0 rpm to 6000 rpm. The testing bearing (ER-12 K deep groove balling bearing from MB) is mounted on the drive end. Fig. 3b–e shows the four health conditions of bearing tested in this work. Vibration and sound signals are widely used in bearing fault diagnosis. Accelerometers mounted on the fault part can capture the vibration signals with higher signal-to-noise ratio (SNR). Sound signals are collected by microphones in a wider frequency bandwidth (20–20000 Hz) than vibration, and are often used in industrial fault diagnosis [21]. As homologous signals, effectively fusing vibration and sound signals can make full use of their complementary characteristic information for bearing fault diagnosis. A mono-axial accelerometer (DYTRAN 3035 B) is mounted on the top of the pedestal, adjacent to the testing bearing for acquiring vibration signals. The sound signal is measured by a microphone (AWA14425), which is installed 20 cm away from the testing bearing. For this experiment, data is collected from different speeds at different speed, with 48000 Hz sample rate. Each test is repeated 3 times to enable statistical robustness.

In this study, linear trend was removed from the vibration and sound data. The result of a linear least-squares fit to data is subtracted from data. Besides, standardization is used to process the data. As a linear transformation, standardization does not lose the information of data, and helps accelerate the calculation of gradient descent. A comparison of the data acquired from four health conditions in time domain and frequency domain is given in Figs. 4 and 5 respectively. It is difficult to judge the health conditions intuitively. The fault frequencies could be extracted from the frequency spectrum but prior knowledge is required for diagnosis. It is expected MF-CNN would get better results while using raw data in frequency domain without other advanced data processing or prior knowledge.

In this section, the MF-CNN model is tested by the experimental data. After pre-processing, signals collected at different working speed are constructed to an input dataset. For each condition of the bearing, data at different working speeds are mixed together to build the training dataset, aiming to verify the performance of the model under non-stationary conditions. The detailed description of the dataset is shown in Table II. Parts of the samples are used to training the model and the rest are used as testing samples.

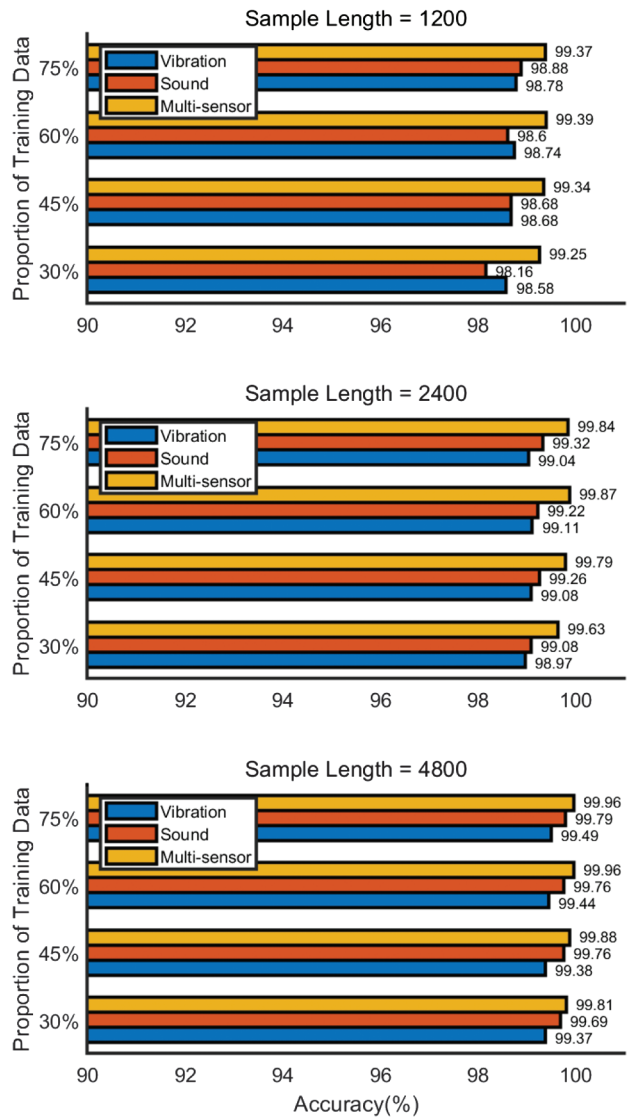


Fig. 6. Testing accuracy of CNN-based methods with different proportion of training data and sample length.

Table III The optimized hyperparameters of classical ML methods

Classical ML Methods	Hyperparameters
kNN	k: 7, weight: distance.
ANN	Layer 1:4800 nodes, Activation: Sigmoid, Layer 2:100 nodes, Activation: Sigmoid, Output Layer: 4 nodes, Activation: Softmax.
SVM	Kernel Type: RBF.

For each experiment, 20 tests are conducted to avoid unexpected contingency in the results, and then calculate the average and standard deviation of these 20 runs.

A comparison among the MF-CNN model and the classical CNN models with single signal as input is constructed and the results are shown in Figure 6. By the way, models with different proportion of training data and different sample length are tested. Accordingly, the

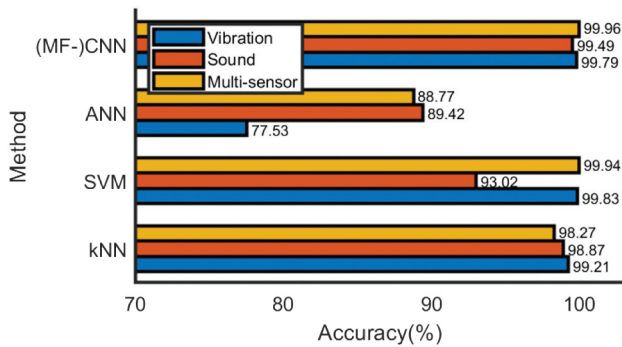


Fig. 7. Testing accuracy of different classification methods.

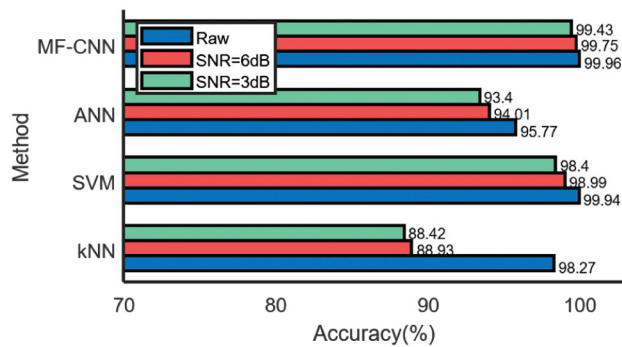


Fig. 8. Testing accuracy of different classification methods with white Gaussian noise disturbing.

MF-CNN model gets better diagnostic results than the typical 1-D CNN model using single signal, which proves that the MF-CNN method is effective for multi-sensor fusion. When the proportion of training data varies from 30% to 75%, the model has good diagnostic ability. With the increase of sample length, the diagnostic accuracy showed a slight growth. However, increasing the sample

length may lead to the increase of training time and GPU memory load. According to the result shown in Fig. 6, 4800 is chosen as the sample length for the subsequent diagnosis.

Classical ML methods, such as k-Nearest Neighbours (kNN), Support vector machine (SVM), Artificial Neural Network (ANN), are used for comparison. The dataset proposed in Table II was adjusted to the suitable format of these method. Hyparameters of each ML method have been carefully optimized and shown in Table III. A comparison of the results of different method with different input data is shown in Fig. 7. The MF-CNN method performs better than the classical ML method, which proves that the method is more effective for bearing fault diagnosis. The achieved results indicate that SVM and the classical CNN model also perform better than other ML methods. It should be noted that, the accuracy of SVM model using both vibration and sound signals is lower than using single sound signal. It proves that some repetitive information in signals may become an obstacle of SVM for classification task. The classical SVM is not a suitable method for multi-sensor signals fusion and fault diagnosis.

Additionally, the measured signals collected in engineering are more complex. Signals acquired from mechanical equipment parts will be coupled together, which brings challenges to fault diagnosis. Therefore, the diagnosis algorithm needs to have the ability to classify the signal with strong noise. In this study, white Gaussian noise with different signal-to-noise ratios (SNR) is manually added to the raw signals, which is conducted by awgn function in MATLAB. 3 dB and 6 dB are two commonly used values and are selected for the test. 3 dB means that the power of the signal is twice that of the noise, and 6 dB means that the power of the signal is four times that of the noise.

The testing results of different classification methods with noise are shown in Fig. 8, which proves that the MF-CNN method proposed has better ability of bearing fault diagnosis with strong noise disturbing. MF-CNN and SVM have higher robustness to noise disturbing. Their diagnostic accuracy is higher than 98%. kNN is sensitive to noise, which reduces accuracy by about 10%. However, it should

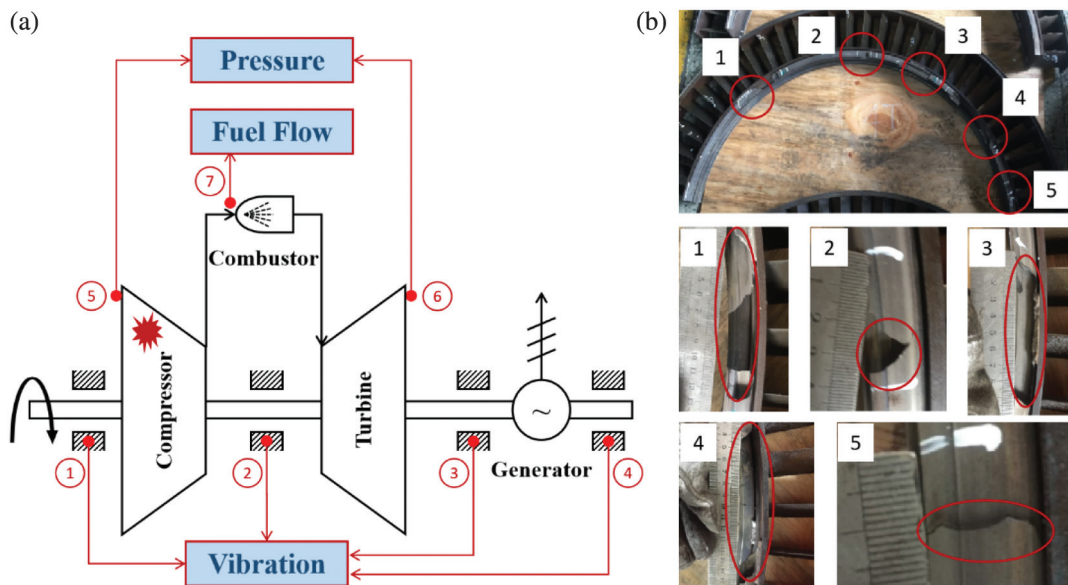


Fig. 9. Detail of the experimental data source. (a) Layout of monitored sensors for gas turbine anomaly detection. (b) Photos of the cracks found during the overhaul.

Table IV Monitored gas turbine sensors for compressor failure

Notion	Variables Description	Variable	Units
1	Rotor vibration on the compressor side	V_{com}	μm
2	Rotor vibration on the turbine side	V_{turb}	μm
3	Generator rotor vibration on the GT side	V_{GTin}	μm
4	Generator rotor vibration on the excitation side	V_{GTex}	μm
5	Compressor inlet pressure	P_{in}	kPa
6	Exhaust pressure	P_{ex}	kPa
7	Fuel flow	M	Nm^3/h

be noted that, while the input with strong noise, the MF-CNN model might face the risk of falling into the local optimal solution or even non-convergence in the process of training. In this experiment, the gross errors of results are selected and deleted. However, in practical engineering, the trained model with high accuracy will be chosen, thus the model is still of practical engineering significance.

B. INDUSTRIAL APPLICAITON ON GAS TURBINE ABNORMAL DETECTION

Experimental validation on gas turbine is conducted in this section. Gas turbine technology is actually considered as a key step toward a low-carbon society [22]. To meet the fluctuation of future energy market, frequent starts and fast load brings challenges to gas turbines [2]. Effective diagnostic methods are essential for the sustainable management of gas turbine plants.

The layout of monitored sensors for gas turbine abnormally detection is shown in Fig. 9a. The pressure signals of the compressor inlet and the turbine exhaust, the vibration signals of four bearings, and the fuel flow signal of combustor are monitored. Their detailed information is listed in Table IV. All data points collected are the peak-to-peak values in 1 minute interval.

In a main overhaul, different cracks were found after inspection on the blades and the seals of the compressor as shown in Fig. 8b. The data provided have been classified as normal and abnormal by the operators, and Fig. 10 displays the signals in different healthy conditions as example [23].

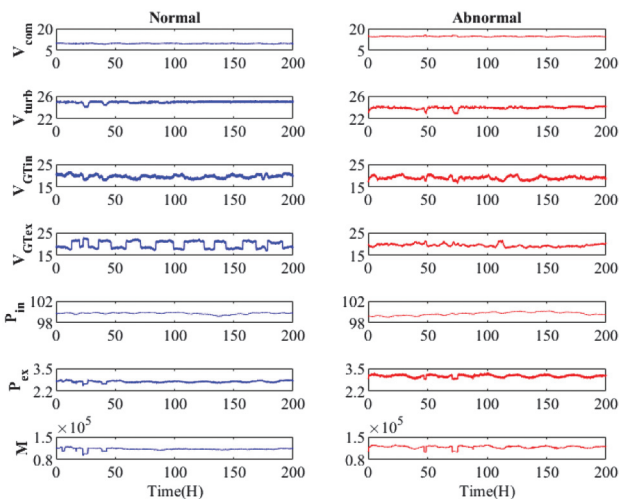


Fig. 10. A comparison among signals collected by 7 sensors.

Table V Testing accuracy of different classification methods for gas turbine abnormal detection

Method	Average accuracy (%)	Standard deviation (%)
kNN	95.29	1.71
SVM	76.62	15.39
ANN	49.67	2.55
MF-CNN	92.06	2.26

The process of dataset establishment is similar to the experiment in section VI(A). MF-CNN and other classical ML methods are tested for comparison and the results are shown in Table V. In this experiment, MF-CNN and kNN method performs well. Although the accuracy of kNN method is higher than MF-CNN, reaching 95.29%, the proposed MF-CNN method has better generalization and stability, combining with the former experiment.

V. CONCLUSIONS

This paper proposed an effective multi-sensor fusion method for fault diagnosis, based on a CNN framework. Signals collected by different sensors are stacked row by row to form a matrix, and the signals are combined together through a convolution operation. In the experiment, some classical ML methods get lower accuracy while using multi-sensor signals. Instead, the proposed MF-CNN model makes better use of the information of multi-sensor signals and fused them in an appropriate way. The model finally reaches a higher accuracy of fault diagnosis.

Moreover, the proposed MF-CNN model has great performance and robustness to different working scenarios, which is verified by serials of experiments. In the two experiments, including laboratory condition, and industrial application condition, the MF-CNN model performs well and successively get high accuracy of 99.96% and 92.06%.

ACKNOWLEDGEMENTS

The authors acknowledge the support from the National Natural Science Foundation of China (Grant No. U1809219) and the Key Research and Development Project of Zhejiang Province (Grant No. 2020C01088).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

References

1. Z. Duan, T. Wu, S. Guo, T. Shao, R. Malekian and Z. Li, "Development and trend of condition monitoring and fault diagnosis of multi-sensors information fusion for rolling bearings: a review," *Int. J. Adv. Manuf. Technol.*, vol. 96, no. 1–4, pp. 803–819, 2018.
2. V. Zaccaria, M. Rahman, I. Aslanidou and K. Kyprianidis, "A review of information fusion methods for gas turbine diagnostic," *Sustainability*, vol. 11, no. 22, p. 6202, 2019.
3. V. Inturi, G. R. Sabareesh, and V. Sharma, "Integrated vibro-acoustic analysis and empirical mode decomposition for fault diagnosis of gears in a wind turbine," *Proc. Struct. Integr.*, vol. 14, pp. 937–944, 2019.
4. M. S. Safizadeh and S. K. Latifi, "Using multi-sensor data fusion for vibration fault diagnosis of rolling element bearings by accelerometer and load cell," *Inform. Fusion*, vol. 18, pp. 1–8, 2014.
5. M. C. Gomes, L. C. Brito, M. Bacci da Silva, and M. A. Viana Duarte, "Tool wear monitoring in micromilling using Support Vector Machine with vibration and sound sensors," *Precis. Eng.*, vol. 67, pp. 137–151, 2021.
6. J. Tang, J. Qiao, Z. Wu, T. Chai, J. Zhang, and W. Yu, "Vibration and acoustic frequency spectra for industrial process modeling using selective fusion multi-condition samples and multi-source features," *Mech. Syst. Signal Process.*, vol. 99, pp. 142–168, 2018.
7. P. Dewallef, C. Romessis, O. Léonard, and K. Mathioudakis, "Combining classification techniques with Kalman filters for aircraft engine diagnostics," *J. Eng. Gas Turbines Power*, vol. 128, no. 2, pp. 281–287, 2006.
8. X. Liu, L. Ma, and J. Mathew, "Machinery fault diagnosis based on fuzzy measure and fuzzy integral data fusion techniques," *Mech. Syst. Signal Process.*, vol. 23, no. 3, pp. 690–700, 2009.
9. W. Jiang, B. Wei, C. Xie, and D. Zhou, "An evidential sensor fusion method in fault diagnosis," *Adv. Mech. Eng.*, vol. 8, no. 3, 2016.
10. H. Zhang and Y. Deng, "Weighted belief function of sensor data fusion in engine fault diagnosis," *Soft Comput.*, vol. 24, no. 3, pp. 2329–2339, 2020.
11. O. Diallo and D. Mavris, "A data analytics approach to failure precursor detection of gas turbine," in *Proceedings of the ASME 2011 Turbo Expo: Turbine Technical Conference and Exposition. Volume 3: Controls, Diagnostics and Instrumentation; Education; Electric Power; Microturbines and Small Turbomachinery; Solar Brayton and Rankine Cycle*, Vancouver, British Columbia, Canada, pp. 305–315, June 2011.
12. Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of machine learning to machine fault diagnosis: A review and roadmap," *Mech. Syst. Signal Process.*, vol. 138, p. 106587, 2020.
13. S. Guo, T. Yang, W. Gao, and C. Zhang, "A novel fault diagnosis method for rotating machinery based on a convolutional neural network," *Sensors*, vol. 18, no. 5, p. 1429, 2018.
14. S. Lu, J. Wang, and Y. Xue, "Study on multi-fractal fault diagnosis based on EMD fusion in hydraulic engineering," *Appl. Thermal Eng.*, vol. 103, pp. 798–806, 2016.
15. L. Zhang, H. Gao, J. Wen, S. Li, and Q. Liu, "A deep learning-based recognition method for degradation monitoring of ball screw with multi-sensor data fusion," *Microelectron. Reliabil.*, vol. 75, pp. 215–222, 2017.
16. M. Azamfar, J. Singh, I. Bravo-Imaz, and J. Lee, "Multi-sensor data fusion for gearbox fault diagnosis using 2-D convolutional neural network and motor current signature analysis," *Mech. Syst. Signal Process.*, vol. 144, p. 106861, 2020.
17. H. Wang, S. Li, L. Song, L. Cui, and P. Wang, "An enhanced intelligent diagnosis method based on multi-sensor image fusion via improved deep learning network," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 6, pp. 2648–2657, 2020.
18. J. Gu, Z. Wang, and J. Kuen, "Recent advances in convolutional neural networks," *Pattern Recogn.*, vol. 77, pp. 354–377, 2018.
19. O. Russakovsky, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2015.
20. I. Sutskever, J. Martens, and G. Dahl, "On the importance of initialization and momentum in deep learning," *Int. Conf. Mach. Learn.*, PMLR, Cambridge MA, pp. 1139–1147, 2013.
21. A. Glowacz, "Acoustic fault analysis of three commutator motors," *Mech. Syst. Signal Process.*, vol. 133, p. 106226, 2019.
22. M. A. M. Ramli, H. R. E. H. Bouchekara, and A. S. Alghamdi, "Efficient energy management in a microgrid with intermittent renewable energy and storage sources," *Sustainability*, vol. 11, no. 14, p. 3839, 2019.
23. X. Gu, S. Yang, Y. Sui, E. Papatheou, A. D. Ball, and F. Gu, "Real-time novelty detection of an industrial gas turbine using performance deviation model and extreme function theory," *Measurement*, vol. 178, p. 109339, 2021.