A New Convolutional Neural Network With Random Forest Method for Hydrogen Sensor Fault Diagnosis

YONGYI SUN[1,2], HONGQUAN ZHANG [1,3], TINGTING ZHAO[3],
ZHIHUI ZOU [3], BIN SHEN[3], AND LIXIN YANG[1]

[1]Key Laboratory of Electronics Engineering, College of Heilongjiang Province, Heilongjiang University, Harbin 150080, China
[2]Department of Information, Liaoning Police Academy, Dalian 116036, China
[3]School of Automation, Harbin Engineering University, Harbin 150001, China

Corresponding author: Hongquan Zhang (1277861501@hrbeu.edu.cn)

ABSTRACT Hydrogen is considered to be a hazardous substance. Hydrogen sensors can be used to detect the concentration of hydrogen and provide an ideal monitoring means for the safe use of hydrogen energy. Hydrogen sensors need to be highly reliable, so fault identification and diagnosis for gas sensors are of vital practical significance. However, traditional machine learning methods for fault diagnosis are based on features extracted by experts, prior knowledge requirements and the sensitivity of system changes. In this study, a new convolutional neural network (CNN) using the random forest (RF) classifier is proposed for hydrogen sensor fault diagnosis. First, the 1-D time-domain data of fault signals are converted into 2-D gray matrix images; this process does not require noise suppression and no signal information is lost. Secondly, the features of the gray matrix images are automatically extracted by using a CNN, which does not rely on expert experience. Dropout and zero-padding are used to optimize the structure of the CNN and reduce overfitting. Random forest, which is robust and has strong generalization ability, is introduced for the classification of gas sensor signal modes, in order to obtain the final diagnostic results. Finally, we design and implement a prototype hydrogen sensor array for experimental verification. The accuracy of fault diagnosis in hydrogen sensors is 100% under noisy environment with the proposed method, which is superior of CNN without RF and other methods. The results show that the proposed CNN with RF method provides a good solution for hydrogen sensor fault diagnosis.

INDEX TERMS Fault diagnosis, hydrogen sensor, convolutional neural network, random forest, feature extraction.

I. INTRODUCTION

Traditional energy sources, such as liquefied petroleum gas, natural gas and coal, are non-renewable resources; therefore, it is crucial that new energy sources are found to replace them. Hydrogen energy is a new energy source for sustainable development, against the background of the conventional energy crisis [1], which has been recognized as a zero-carbon energy source. In the 21st century, progress has been

The associate editor coordinating the review of this manuscript and approving it for publication was Aijun Yang .

made in many aspects in the field of hydrogen energy. Many advanced countries have formulated plans for the development of hydrogen energy [2], [3].

Hydrogen—a highly reactive molecule—is considered to be a hazardous substance. It is flammable and explosive [4], so it is particularly important to monitor hydrogen leakage for safety purposes. A hydrogen sensor, which is designed to monitor the concentration of hydrogen, is a necessary device for the safe use of hydrogen energy [5]. Once the concentration exceeds the normal range, its alarm will sound immediately. However, due to the influence of environmental

factors, hydrogen sensors are prone to failure; once one fails in application, it will lose its hydrogen safety detection function, which may lead to combustion and explosion. Therefore, it is vital to detect and identify faults in hydrogen sensors.

Semiconductor gas sensors based on SnO_2 sensitive materials have been widely used in practical applications and are a mature technology [6]. However, in terms of the effectiveness of their long-term use, the output value of the gas sensor is not only related to the concentration of the measured gas but is also affected by environmental factors, such as dust, humidity, temperature and air pressure, as well as degradation in the chemical characteristics of the sensor materials (e.g., heating of the wire or oxidation). All of these factors lead to the parameter drift of a gas sensor and impair its effectiveness. Therefore, fault diagnosis of gas sensors has become an important issue to many researchers [7]–[11].

Generally, fault diagnosis methods for sensors can be divided into four categories: knowledge-based, model-based, data-driven and hybrid/active [12]. As the data-driven method is suitable for the analysis of complex signal systems, it has been widely used in fault diagnosis, and an increasing number of engineers and researchers are implementing this method in their work [13].

Machine learning (ML) methods are an important aspect of the data-driven approach, which have already been applied to fault diagnosis; for example, the k-nearest neighbors algorithm (KNN) [11], principal component analysis [14], extreme learning machine (ELM) [15], wavelet transform [16], support vector machine (SVM) [17], learning vector quantization (LVQ) [18], [19], back propagation (BP) neural network [20], gray forecasting [21] and non-negative matrix factorization [22] models have all been used for fault diagnosis.

The feature extraction process in these methods is unable to generate discriminative features of raw data, as well as consuming a lot of time and energy. Thus, the degree of automation is greatly reduced. The final result can also be impacted by the extraction process [23]. If a system is particularly complex, choosing an appropriate feature function requires considerable expertise and a deep mathematical basis. Thus, expert experience will directly affect the final results.

Recently, deep learning (DL), as an advanced technology, has been able to overcome the above shortcomings. It can automatically learn abstract features of the original data and classify them effectively, avoiding the shortcoming of requiring handcrafted features designed by engineers [24]. Therefore, many DL methods have been gradually introduced into fault diagnosis, such as sparse filtering [25], deep belief networks [26], [27] and sparse autoencoders [28]. Convolutional neural network (CNN) is one of the most effective types of DL. Turker Ince *et al.* used a 1-D CNN to detect motor faults [29]; Long Wen *et al.* used a CNN for fault diagnosis [30]; and Qinyu Jiang *et al.* used a deep CNN for bearing fault diagnosis in a noisy environment [31]. Min Xia *et al.* used a CNN-based approach for fault diagnosis of rotating machinery [32]. An ensemble CNN model is proposed for

bearing fault diagnosis by Yang Liu et al [33]. Also, a novel sensor data-driven fault diagnosis method is proposed based on CNN [34]. Compared to traditional ML methods, a CNN has achieved better results, but its application in gas sensor fault diagnosis is still in a developmental stage.

Classification is an essential part of the process of fault diagnosis, therefore the choice of classifier is very important. Ensemble learning, also called multi-classifier system and committee-based learning, uses multiple weak classifiers to form a strong classifier, the classification results are obtained according to the majority voting principle. Such a process makes it perform better for complex data and obtain better classification accuracy and generalization performance. The data of fault diagnosis has the characteristics of large scale, multi-scale and autocorrelation, thus the classifier based on ensemble learning could obtain better performance in fault diagnosis [35], [36].

Random forest (RF) is an ensemble learning method which classifies using a voting model. Compared to other ML methods, RF has the characteristics of low complexity, fast computing speed, high accuracy rate, insensitive to parameters, no need for feature normalization, less over fitting, etc. [37], [38]. Importantly, RF is more robust with respect to noise [37]. Therefore, it is more suitable to use RF when respect to a large number of data with reasonable features, especially under noisy environment. The literature [38], [39] has demonstrated that sensors fault diagnosis based on RF is feasible.

It has been reported that the advantage of CNN's ability in feature extraction combined with the good performance of classification of RF has been adopted for image classification [40], solar photovoltaic array detection [41], internet intrusion detection [42], scene categorization [43], facial expression recognition [44], tree species classification [45], ship identification on satellite image [46]. Recently, a novel bearing fault diagnosis method based on CNN and RF has been proposed, and experimental results indicate that the proposed method achieves high accuracy in bearing fault diagnosis under complex operational conditions and is superior to traditional methods and standard deep learning methods [47]. However, the novel method combined CNN and RF is rarely used for gas sensor fault diagnosis till now.

In this article, a method for hydrogen sensor fault diagnosis using a CNN with RF (CNN-RF) is proposed to automatically capture features of the gas sensor signal and improve upon the performance of conventional methods. The main contributions of the paper are as follows.

1) A method for transforming raw fault data into gray matrix images is proposed to process the sensor fault signal directly, which does not require expert experience.

2) In order to reduce the overfitting phenomenon, the structure of the CNN is optimized by dropout and zero-padding.

3) The sensor signal features captured by the CNN are input into the RF classifier to diagnose the fault mode

of a hydrogen sensor. The proposed method is verified by a self-made experimental system. The experimental results show that the accuracy of the CNN-RF method is higher than the accuracy of the CNN alone and other methods.

The remainder of this paper is organized as follows: The second section introduces the theoretical fundamentals. In the third section, a novel model based on CNN-RF for hydrogen sensor fault diagnosis is introduced. The fourth section verifies the effectiveness of the proposed method through experiments. The fifth section provides a summary and discussion.

II. THEORETICAL FUNDAMENTALS

Convolutional neural network was proposed in the late 1980s for processing data in the form of multiple arrays [48]. Firstly, in a CNN, each neuron in a feature map is sparsely connected to a small group of neurons in the previous layer, which is different from the connections in an artificial neural network (ANN). A CNN mainly comprises a convolutional layer, a pooling layer and a fully connected (FC) layer. As CNN was inspired by the concept of simple and complex cells in the visual cortex in the brain, it has been widely used in computer vision and image classification [48].
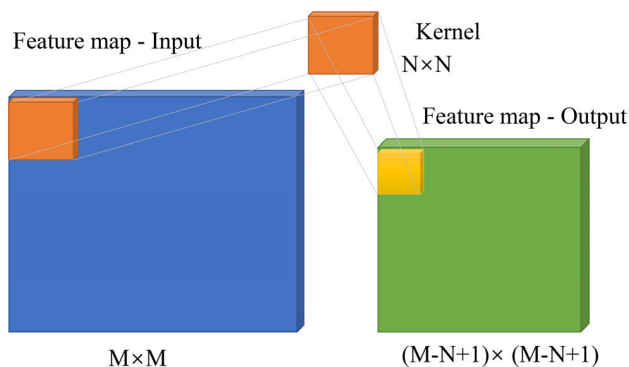

FIGURE 1. Convolutional layer.

A. CONVOLUTIONAL LAYER

The purpose of convolutional layers is to extract different input features. Each feature map is composed of rectangular neurons. Neurons in the same feature map share weights, which are called the convolutional kernel. Convolutional kernels are usually initialized in the form of random matrices. A convolutional layer is shown in Figure 1. The direct benefit of using shared weights (i.e., convolutional kernels) is to reduce the connection between layers of the network while reducing the risk of overfitting [49], [50].

In a convolutional layer, assume that there are K filters and M is the input. Generally, the output feature maps of the lth layer are calculated as follows [50]:

$$x_j^l = f \left(\sum_{i=1,\dots,M} x_i^{l-1} \times k_{ij}^l + b_j^l \right), \quad j = 1, \dots, N, \quad (1)$$

where f and x are the activation function and convolutional operations, respectively; b_j^l denotes the bias corresponding to

jth filter, x_i^{l-1} denotes the ith input map, x_j^l denotes the jth output map, and k_{ij}^l denotes the kernel of the jth filter, which is connected to the ith input map.

B. POOLING LAYER

Pooling layers are a form of downsampling. The function of the pooling layer is that it can merge similar features in a local position to make the detection more reliable [26]. In a pooling layer, assume that there are M input feature maps and M output feature maps. Generally, the output feature maps of the lth layer are calculated as follows [49]:

$$x_j^l = f \left(\beta_j^l \, down \left(x_j^{l-1} \right) + b_j^l \right), \quad j = 1, \dots, M, \quad (2)$$

where f is the activation function, *down* is the sub-sampling function, β_j^l and b_j^l are the multiplicative bias and the additive bias corresponding to the jth filter, respectively, x_j^l is the jth output map and x_j^{l-1} is the jth input map.
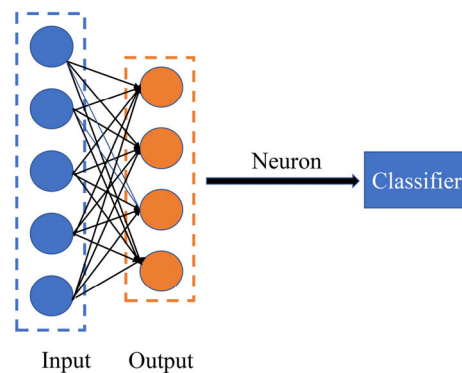

FIGURE 2. Fully connected layer.

C. FULLY CONNECTED LAYER

In a CNN structure, after several convolutional layers and pooling layers, one or more FC layers are connected. The FC layer model is shown in Figure 2. Assuming that the length of the input is M and the total length of the output vector is N, then the output vector of the lth layer is calculated as follows [51]:

$$x_j^l = \left(\sum_{i=1,\dots,M} x_i^{l-1} \times w_{ij}^l + b_j^l \right), \quad j = 1, \dots, N, \quad (3)$$

where f is the activation function, x_i^{l-1} is the jth input value, x_j^l is the jth output value, b_j^l represents the bias corresponding to the jth output value and w_{ij}^l is the weight of the jth output value, which is connected to the ith input value.

There are many famous CNN models, such as GoogLeNet [52], LeNet-5 [53], AlexNet [54], and Network in Network [55]. In this paper, the classical release of CNN, LeNet-5, which has been applied to handwritten character recognition, is adopted to solve the gray matrix image classification task of fault diagnosis. It has two alternating convolutional and pooling layers with a two-layer FC ANN.

III. PROPOSED MODEL FOR HYDROGEN SENSOR FAULT DIAGNOSIS

In this section, a new model is proposed for hydrogen sensor fault diagnosis. Firstly, a method for image conversion from raw data to a gray matrix image is proposed for processing of the sensor signal. Secondly, two strategies—dropout and zero-padding—are used to optimize the CNN structure. Finally, the novel CNN-RF method is proposed.

A. GRAY MATRIX IMAGE CONVERSION METHOD

Data pre-processing of the sensor signal is the first step in gas sensor fault diagnosis; the quality of data processing will directly affect the accuracy of the diagnosis. Traditional gas sensor fault signal processing methods mostly rely on expert experience to extract features from raw data and cannot directly handle raw signals [56]. Extracting features is not only exhausting work but also plays a key role in the results. In this study, a method for transforming the raw gas sensor signal into gray matrix images is proposed. The main aspects of the method are as follows.

Algorithm 1 Algorithm of Gray Matrix Image Conversion

Input: The measured raw data sequence X_i of each sensor in the sensor array.
Output: The image of gray matrix conversion.
1: Define: temp = raw data sequence.
2: Convert the temp to a value between 0 and 1.
3: get the image_temp when image_temp = uint8 (temp × 255).
4: **for** j = 1: M then
5: get the image value: image (j, 1: N) = image_temp ((1 + (j − 1) × N): (N × j)).
6: end
7: **end for**

The method of gray matrix image conversion is sketched in Algorithm 1. In this method, the raw signal of the gas sensor is first converted into a value between 0 and 1, followed by conversion into a gray matrix image of dimension M × N using uint8 encoding technology, where M is the width of the image and N is the height of the image. The gray matrix image conversion method is shown in Figure 3.

The advantage of this method is that it provides a representation for exploration of the 2-D features of the original sensor signal. It can retain the original features of the data as much as possible and does not depend on expert experience or artificial feature extraction.

In this paper, as each 1-D fault signal sample consisted of 2000 data elements, the gray matrix image size was set to 50 × 40 pixels.

B. DROPOUT

In CNN training, the problem of overfitting is often encountered. Reducing the interaction of feature detectors in a neural network can prevent overfitting, in order to improve its performance [57]. Dropout can be used as a strategy for
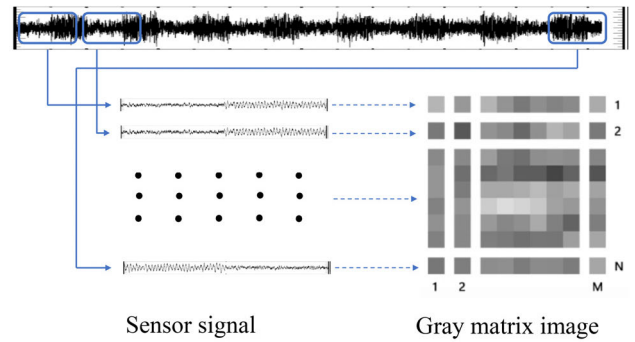

FIGURE 3. Gray matrix image conversion method.

training deep neural networks where, in each training batch, the overfitting phenomenon can be significantly reduced by ignoring half of the feature detectors; that is, the weights of half of the hidden layer nodes are set to 0 during dropout.

In this study, dropout is used to effectively prevent overfitting. We set the retaining probability to $p = 0.5$ for dropout and, so, the output of each neuron in every layer was zero with probability 0.5. The dropout neural net model is shown in Figure 4. This strategy can reduce interactions between feature detectors.
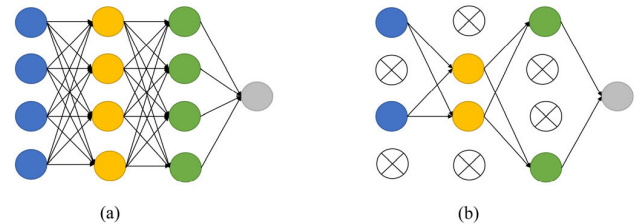

FIGURE 4. Dropout neural net model (a) standard neural net and (b) with dropout neural net.

C. ZERO-PADDING

In processing image information using a CNN, a majority of the edge pixels in the input image are only operated on by the convolutional kernel once, whereas pixels in the middle of the image will be scanned many times. This reduces the reference degree of boundary information, to a certain extent. On the other hand, after using zero-padding, the new boundary has an effect on a certain part of the actual processing. This problem can be solved, to a degree. At the same time, input images of different sizes can be complemented such that they are same size. Suppose the input size is (H, W), the filter size is (FH, FW), the output size is (OPH, OPW), the padding length is P, and the stride size is S. Then, the output size formula is as follows:

$$OPH = \frac{H + 2P - FH}{S} + 1 \qquad (4)$$

$$OPW = \frac{W + 2P - FW}{S} + 1 \qquad (5)$$

Therefore, in this study, zero-padding is used to control the feature dimension.

D. RF CLASSIFIER

The selection of a classifier plays a key role in the classification results. It is difficult for the traditional CNN, which is based on the softmax classifier, to achieve the best generalization capability, due to the local minimum, vanishing gradient, and overfitting problems in the training process.

The random forest algorithm is an ensemble learning method based on a decision tree. Suppose there are N training sets, and each tree randomly selects N training samples from them as a sub-training set. If there are M features, select m (m < M) features, and then select the optimal feature from each split. In this way, each tree can obtain training results according to different sub-training sets, and sampling with return can also ensure the "integrity" of the training results. Each tree of the input sample is judged separately and the final classification is determined according to the voting results; that is, the results of several weak classifiers are combined to form a strong classifier. The proposed model makes use of the RF as an initial "mock test" in the algorithm, and thus has good application in large data sets and for input samples with high-dimensional features. The RF model is shown in Figure 5.
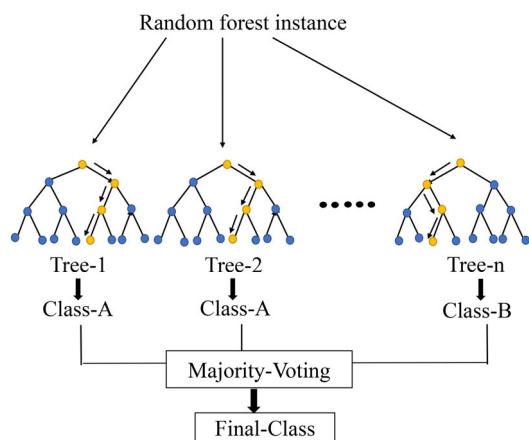

FIGURE 5. Random forest (RF) model.

Therefore, in this study, RF is used as a classifier. When only a few parameters need to be adjusted, its robustness against noise disturbance is enhanced, the generalization ability and classification effect of the model are improved and overfitting of the model is reduced.

E. CNN-RF METHOD FOR FAULT DIAGNOSIS

This paper proposes the CNN-RF method, which contains four types of layers in its structure: namely, a convolutional layer, a pooling layer, a fully connected layer and a RF classifier layer. It also uses two strategies—dropout and zero-padding—to prevent overfitting and enhance its performance. The CNN-RF method is used to process gray matrix images for fault diagnosis of hydrogen sensors. The samples are labeled by fault type, according to the sensor signal records. Then, the samples were converted into gray matrix images. Layers with dropout and zero-padding were added to deal

with the overfitting problem. The sensor fault type features are captured by the convolutional, pooling, and fully connected layers. Then, the features are input into the RF classifier to obtain the fault diagnosis results, as it has demonstrated good performance in fault mode classification of hydrogen sensors. The overall structure of the proposed model based on the CNN-RF method introduced in this paper is shown in Figure 6.

IV. EXPERIMENT AND VALIDATION OF THE PROPOSED METHOD

A. EXPERIMENTAL SETUP

The data used for verification of the CNN-RF method were obtained through our experimental system. A system diagram of hydrogen sensor arrays is shown in Figure 7. The experimental system included a standard hydrogen concentration cylinder, a standard air cylinder, a gas molecular flowmeter, a gas mixer, a two-way regulated power supply, a data collector, a constant temperature and constant humidity box, a computer system, a sensor array chamber, a SnO_2 sensor array and a sensor model (MQ-8). We used a six-sensor, commercially available MQ-8 gas sensor array, screen-printed in our experiments, as shown in Figure 8. The MQ-8 gas sensor cylinder core structure is shown in Figure 9.

When the test system works, the standard gas in the cylinder enters the gas molecule flowmeter through the pressure-reducing valve, which controls the flow rate of the air molecule flowmeter and the standard hydrogen molecule flowmeter in order to obtain the target hydrogen concentration in the gas mixer. After the hydrogen concentration in the gas mixer is homogenized, it flows into the gas chamber of the sensor array through the pipeline and is loaded onto the sensor array. The sensor array detects the concentration of hydrogen gas after the working and heater voltages are provided by the two-way regulated power supply, respectively. The detection signal of the sensor array is picked up by the data collector and transmitted to the computer system through a 232 data bus interface. The structure of the sensor signal pickup circuit is shown in Figure 10. The program was run on a 2.8 GHz Intel CPU with 8 GB RAM running Windows 10. A photographic image of the experimental setup of the hydrogen sensor array is shown in Figure 11.

According to long-term practical experience and related literature reports [7], [58], [59], six gas sensor fault types were selected: impact fault, stuck fault, heating wire disconnection (HWD) fault, bias fault, exfoliation of sensitive body (ESB) fault and false welding of sensitive body (FWSB) fault. Of them, the structure of HWD fault and ESB fault is shown in Figure 12, which could provide a good understanding of hydrogen sensor's fault structure. Through the experiments, we obtained MQ-8 gas sensor signals of seven modes under normal environment (i.e., the six fault types and without fault), as shown in Figure 13.

However, hydrogen sensors perform under noisy environment is inevitable in real world industrial applications.
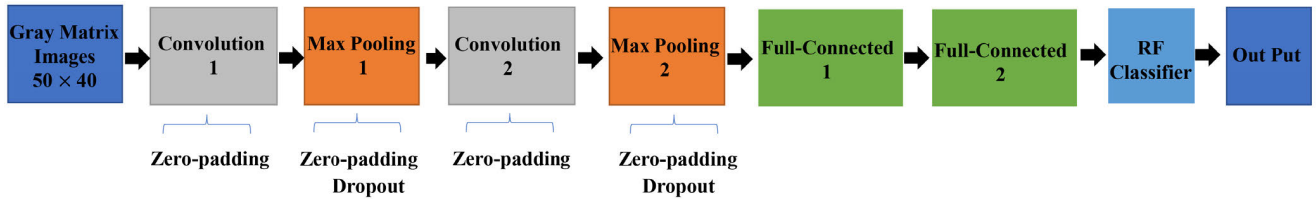
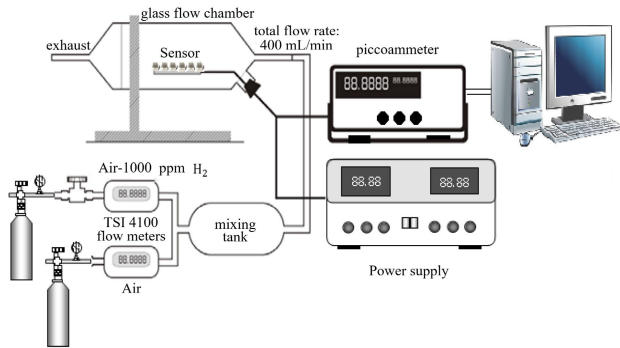**FIGURE 6.** Proposed model based on convolutional neural network-random forest (CNN-RF) in this paper.
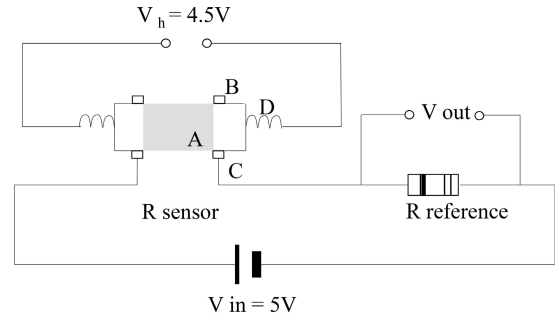

FIGURE 7. System diagram of hydrogen sensor arrays.

FIGURE 10. MQ-8 gas sensor signal picking circuit diagram (A) SnO_2 film, (B) Au electrode, (C) Pt wire and (D) Ni-Cr alloy resistor.

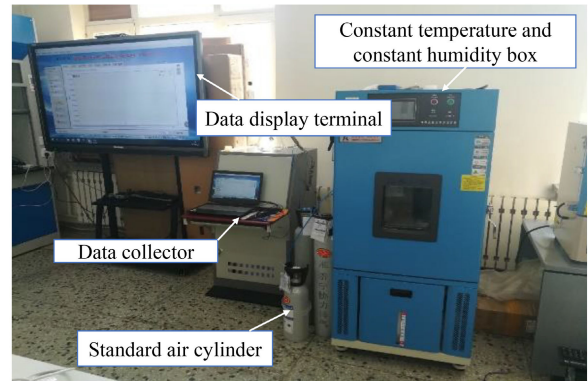**FIGURE 8.** A photograph of the MQ-8 gas sensor array used in our experiments.

FIGURE 11. Experimental setup of MQ-8 gas sensor array.
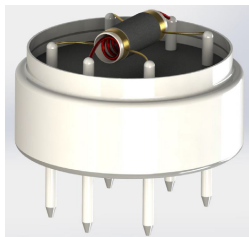
FIGURE 9. MQ-8 gas sensor cylinder core structure.
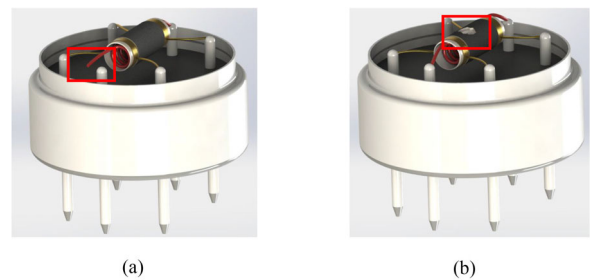
FIGURE 12. The structure of HWD fault and ESB fault (a) HWD fault and (b) ESB fault.

Since the noise varies a lot, and we can't get all the labeled training samples under different noisy environment. Accordingly, additive white Gaussian noise is added to the original signals to composite signals [24]. MQ-8 gas sensor signals of seven modes under noisy environment (i.e., the six fault types and without fault) are shown in Figure 14.

The CNN-RF method requires a lot of data for training samples and test samples, and sensor fault signals are difficult to obtain in large quantities; therefore, based on the acquired (normal and fault) signals, data simulation was carried out

FIGURE 13. Seven signal modes of MQ-8 gas sensor under normal environment.

Normal signal

The signal including impact fault The signal including stuck fault

The signal including HWD fault The signal including bias fault

The signal including ESB fault The signal including FWSB fault

FIGURE 14. Seven signal modes of MQ-8 gas sensor under noisy environment.
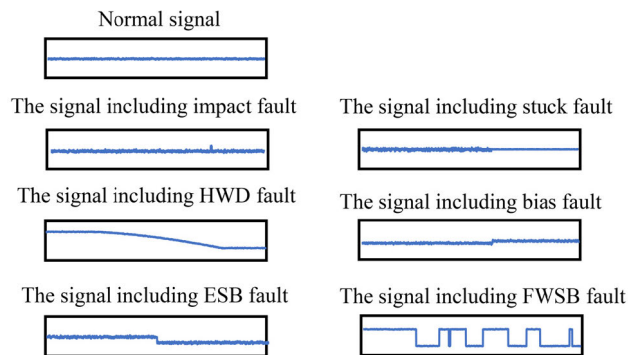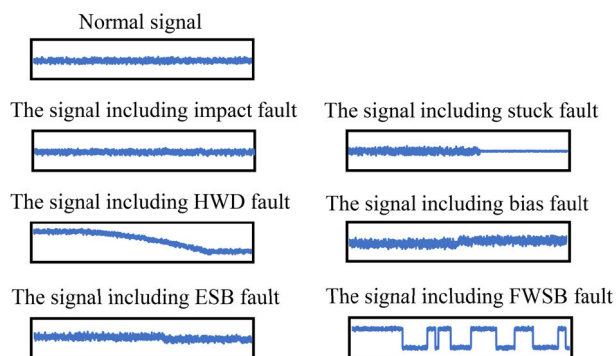
Normal signal

The signal including impact fault The signal including stuck fault

The signal including HWD fault The signal including bias fault

The signal including ESB fault The signal including FWSB fault

Normal signal Impact fault Stuck fault HWD fault

Bias fault ESB fault FWSB fault

FIGURE 15. Converted gray matrix images of the gas sensor signals of seven modes under noisy environment.

converted into gray matrix images, the size of each being 50×40 pixels (as shown in Figure 15). Then, the images were input into the proposed method for training.

The proposed method was trained for 1000 iterations. The relevant parameters for each layer of CNN-RF are listed in Table 2. The training accuracy reached 100% and the training loss closed to 0 after 300 iterations as is shown in Figure 16. Furthermore, all values remained stable after 300 iterations. Thus, we obtained a well-trained model with 100% accuracy through training of the CNN-RF.

TABLE 2. Parameters of CNN-RF.

Layer No.	Layer type	Kernel size /stride	Kernel channel size	Drop-out	Zero-padding
1	Convolution 1	$3 \times 3 / 1 \times 1$	16	No	Yes
2	Max Pooling 1	$3 \times 3 / 2 \times 2$	16	Yes	Yes
3	Convolution 2	$3 \times 3 / 1 \times 1$	16	No	Yes
4	Max Pooling 2	$3 \times 3 / 2 \times 2$	16	Yes	Yes
5	FC 1	2080	1	No	No
6	FC 2	128	1	No	No
7	RF	7	1	No	No

to increase the amount of data. The simulation samples were acquired by overlapping a sensor fault signal onto the normal signal acquired by the sensor array.

B. VALIDATION OF CNN-RF METHOD
In this section, the results of CNN-RF training and inference are provided, in order to validate the performance of the proposed model in hydrogen sensor fault diagnosis.

TABLE 1. Sample number of seven modes of sensor signal.

Models	Fault description	Training samples	Testing samples
1	Normal signal	100	37
2	Impact fault	100	37
3	Stuck fault	100	37
4	HWD fault	100	37
5	Bias fault	100	37
6	ESB fault	100	37
7	FWSB fault	100	37

1) CNN-RF TRAINING
As shown in Table 1, 137 sets of experimental samples were obtained for each sensor signal mode, of which 100 were training samples and 37 were test samples. There was no repetition between the training samples and test samples. The seven sensor signal modes under noisy environment were first
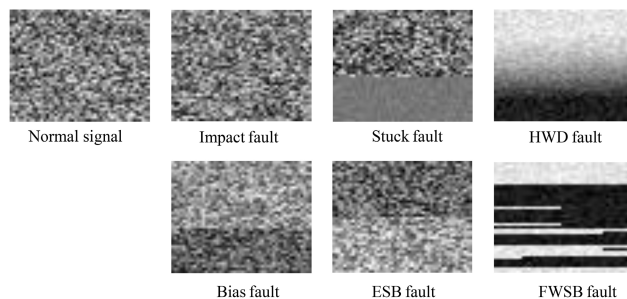
2) CNN-RF INFERENCE
Inference for hydrogen sensor fault diagnosis was realized by the new method, according to the experimental data. To obtain better results, the experiment was repeated six times, and the final diagnosis results of CNN-RF and CNN were compared under normal and noisy environment (in terms of accuracy). The mean accuracy results of CNN-RF and CNN are listed in Table 3. The diagnosis results for the seven hydrogen sensor signal modes under noisy environment using CNN-RF are shown in Figure 17, where the matching degree between the predicted and actual type is 100% in each row.

In order to further verify the advantage of RF as the classifier involved in the method of CNN-RF we adopted, we replaced a RF classifier by a KNN (CNN-KNN), a SVM (CNN-SVM), a BP (CNN-BP) classifiers and compared the mean accuracy of all the methods above under noisy
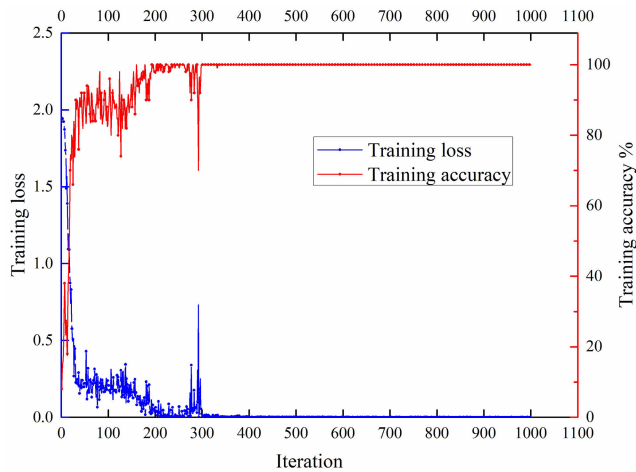
FIGURE 16. Proposed method training accuracy and loss under noisy environment.

TABLE 3. The mean accuracy of CNN-RF and CNN under normal and noisy environment repeated 6 times.

Method	Normal environment		Noisy environment	
	CNN-RF	CNN	CNN-RF	CNN
Max	100%	100%	100%	99.61%
Min	100%	99.23%	100%	98.84%
Mean	100%	99.74%	100%	98.91%

FIGURE 17. Test results in the diagnosis of the seven hydrogen sensor signal modes using CNN-RF under noisy environment.

TABLE 4. The mean accuracy of CNN-RF, CNN-KNN, CNN-SVM and CNN-BP under noisy environment repeated 6 times.

Method	CNN-RF	CNN-KNN	CNN-SVM	CNN-BP
Max	100%	99.61%	99.61%	99.23%
Min	100%	98.84%	98.84%	98.84%
Mean	100%	99.16%	99.29%	98.91%

environment repeated 6 times. The experimental results show that the CNN-RF model has a stronger robustness and higher generalization ability for the non-linear problem of hydrogen sensor fault diagnosis, all the results are listed in Table 4.

To evaluate the performance of the CNN-RF method, other traditional methods were selected for a comparison of prediction accuracy. The selected methods were KNN [11], ELM [15], SVM [17], LVQ [18], [19], BP [20], RF [39], and CNN-RF, the last of which (i.e., the proposed model) had higher accuracy than the other methods under

TABLE 5. Diagnosis accuracy based on seven different methods under noisy environment.

Methods	Accuracy (%)
KNN	71.81%
ELM	70.27%
SVM	84.56%
LVQ	66.41%
BP	74.52%
RF	96.91%
CNN-RF	100%

noisy environment. The comparison results, in terms of sensor fault diagnosis accuracy, are shown in Table 5.

V. CONCLUSION

In this study, we presented a novel method, CNN-RF, for the fault diagnosis of hydrogen sensors. The proposed method is able to fuse the three major blocks of traditional fault-detection approaches into a single learning body—feature extraction, feature selection and classification—without requiring expert intervention. The main contributions of this study are the development of a raw sensor signal-to-gray matrix images conversion method, which changes the default image size of LeNet-5 from 32×32 pixels to 50×40 pixels, according to the length of the hydrogen sensor signal data. Through dropout and zero-padding, the structure of a CNN with a RF classifier is optimized, which ensures that the training model structure of the proposed method has better generalization ability and robustness, compared to the traditional CNN method, for the fault diagnosis of hydrogen sensors. The experimental results show that the proposed method can learn features effectively and achieve convincing detection results for a hydrogen sensor with seven modes. The proposed method achieved a prediction accuracy of 100% on the seven modes studied, outperforming the CNN alone and other methods. The proposed method can also be applied to other gas sensor fault diagnosis.

Some limitations of the new method are the following: optimization of the CNN parameters requires debugging on a case-by-case basis, and the effectiveness of the method has only been verified in this experiment. Based on the above limitations, we plan to consider how optimization techniques can be used to adjust the CNN parameters in fault detection for process monitoring and to apply the method to a larger range of sensor fault diagnosis scenarios in future work.

REFERENCES

[1] P. Nikolaidis and A. Poullikkas, "A comparative overview of hydrogen production processes," *Renew. Sustain. Energy Rev.*, vol. 67, pp. 597–611, Jan. 2017.

[2] S. G. Chalk and J. F. Miller, "Key challenges and recent progress in batteries, fuel cells, and hydrogen storage for clean energy systems," *J. Power Sour.*, vol. 159, no. 1, pp. 73–80, Sep. 2006.

[3] T. He, P. Pachfule, H. Wu, Q. Xu, and P. Chen, "Hydrogen carriers," *Nature Rev. Mater.*, vol. 1, no. 12, 2016, Art. no. 16059.

[4] H. Lee, J.-R. Sohn, S.-H. Byeon, S. Yoon, and K. Moon, "Alternative risk assessment for dangerous chemicals in South Korea regulation: Comparing three modeling programs," *Int. J. Environ. Res. Public Health*, vol. 15, no. 8, p. 1600, 2018.

[5] Y. Hu, J. Lei, Z. Wang, S. Yang, X. Luo, G. Zhang, W. Chen, and H. Gu, "Rapid response hydrogen sensor based on nanoporous Pd thin films," *Int. J. Hydrogen Energy*, vol. 41, no. 25, pp. 10986–10990, Jul. 2016.

[6] G. Fedorenko, L. Oleksenko, N. Maksymovych, G. Skolyar, and O. Ripko, "Semiconductor gas sensors based on Pd/SnO_2 nanomaterials for methane detection in air," *Nanosc. Res. Lett.*, vol. 12, no. 1, p. 329, Dec. 2017.

[7] Y.-S. Chen, Y.-H. Xu, J.-L. Yang, Z. Shi, S.-D. Jiang, and Q. Wang, "Fault detection, isolation, and diagnosis of status self-validating gas sensor arrays," *Rev. Sci. Instrum.*, vol. 87, no. 4, Apr. 2016, Art. no. 045001.

[8] S. U. Jan, Y.-D. Lee, J. Shin, and I. Koo, "Sensor fault classification based on support vector machine and statistical time-domain features," *IEEE Access*, vol. 5, pp. 8682–8690, 2017.

[9] M. Pardo, G. Faglia, G. Sberveglieri, M. Corte, F. Masulli, and M. Riani, "Monitoring reliability of sensors in an array by neural networks," *Sens. Actuators B, Chem.*, vol. 67, nos. 1–2, pp. 128–133, Aug. 2000.

[10] J. Yang, Y. Chen, and L. Zhang, "An efficient approach for fault detection, isolation, and data recovery of self-validating multifunctional sensors," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 3, pp. 543–558, Mar. 2017.

[11] J. Yang, Z. Sun, and Y. Chen, "Fault detection using the clustering-kNN rule for gas sensor arrays," *Sensors*, vol. 16, no. 12, p. 2069, 2016.

[12] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques—Part I: Fault diagnosis with model-based and signal-based approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3757–3767, Jun. 2015.

[13] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *Int. J. Inf. Manage.*, vol. 35, no. 2, pp. 137–144, Apr. 2015.

[14] E. P. Tao, W. H. Shen, T. L. Liu, and X. Q. Chen, "Fault diagnosis based on PCA for sensors of laboratorial wastewater treatment process," *Chemometric Intell. Lab. Syst.*, vol. 128, pp. 49–55, Oct. 2013.

[15] J. Lu, J. Huang, and F. Lu, "Sensor fault diagnosis for aero engine based on online sequential extreme learning machine with memory principle," *Energies*, vol. 10, no. 1, p. 39, 2017.

[16] R. Yan and R. X. Gao, "Hilbert–Huang transform-based vibration signal analysis for machine health monitoring," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 6, pp. 2320–2329, Dec. 2006.

[17] F. Deng, S. Guo, R. Zhou, and J. Chen, "Sensor multifault diagnosis with improved support vector machines," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 2, pp. 1053–1063, Apr. 2017.

[18] N. Wan, W. Zhang, and K. Li, "Fault identification of high frequency current ripple of fuel cell based on learning vector quantization neural network," *DEStech Trans. Eng. Technol. Res.*, 2017.

[19] A. M. Bassiuny, X. Li, and R. Du, "Fault diagnosis of stamping process based on empirical mode decomposition and learning vector quantization," *Int. J. Mach. Tools Manuf.*, vol. 47, no. 15, pp. 2298–2306, Dec. 2007.

[20] Q. Liang, H. Han, X. Cui, H. Qing, and Y. Fan, "Comparative study of probabilistic neural network and back propagation network for fault diagnosis of refrigeration systems," *Sci. Technol. Built Environ.*, vol. 24, no. 4, pp. 448–457, Apr. 2018.

[21] Y. Chen, J. Yang, Y. Xu, S. Jiang, X. Liu, and Q. Wang, "Status self-validation of sensor arrays using gray forecasting model and bootstrap method," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 7, pp. 1626–1640, Jul. 2016.

[22] J.-L. Yang, Y.-S. Chen, L.-L. Zhang, and Z. Sun, "Fault detection, isolation, and diagnosis of self-validating multifunctional sensors," *Rev. Sci. Instrum.*, vol. 87, no. 6, Jun. 2016, Art. no. 065004.

[23] N. Zeng, Z. Wang, H. Zhang, W. Liu, and F. E. Alsaadi, "Deep belief networks for quantitative analysis of a gold immunochromatographic strip," *Cognit. Comput.*, vol. 8, no. 4, pp. 684–692, Aug. 2016.

[24] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mech. Syst. Signal Process.*, vol. 100, pp. 439–453, Feb. 2018.

[25] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3137–3147, May 2016.

[26] H. Shao, H. Jiang, X. Zhang, and M. Niu, "Rolling bearing fault diagnosis using an optimization deep belief network," *Meas. Sci. Technol.*, vol. 26, no. 11, Nov. 2015, Art. no. 115002.

[27] W. He, P.-L. Qiao, Z.-J. Zhou, G.-Y. Hu, Z.-C. Feng, and H. Wei, "A new belief-rule-based method for fault diagnosis of wireless sensor network," *IEEE Access*, vol. 6, pp. 9404–9419, 2018.

[28] L. Wen, L. Gao, and X. Li, "A new deep transfer learning based on sparse auto-encoder for fault diagnosis," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 1, pp. 136–144, Jan. 2019.

[29] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016.

[30] L. Wen, X. Li, L. Gao, and Y. Zhang, "A new convolutional neural network-based data-driven fault diagnosis method," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5990–5998, Jul. 2018.

[31] Q. Jiang, F. Chang, and B. Sheng, "Bearing fault classification based on convolutional neural network in noise environment," *IEEE Access*, vol. 7, pp. 69795–69807, 2019.

[32] M. Xia, T. Li, L. Xu, L. Liu, and C. W. de Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 1, pp. 101–110, Feb. 2018.

[33] Y. Liu, X. Yan, C.-A. Zhang, and W. Liu, "An ensemble convolutional neural networks for bearing fault diagnosis using multi-sensor data," *Sensors*, vol. 19, no. 23, p. 5300, 2019.

[34] G. Li, C. Deng, J. Wu, X. Xu, X. Shao, and Y. Wang, "Sensor data driven bearing fault diagnosis based on deep convolutional neural networks and S-transform," *Sensors*, vol. 19, no. 12, p. 2750, 2019.

[35] S. Ma and F. Chu, "Ensemble deep learning-based fault diagnosis of rotor bearing systems," *Comput. Ind.*, vol. 105, pp. 143–152, Feb. 2019.

[36] Z. Zhang, H. Han, X. Cui, and Y. Fan, "Novel application of multi-model ensemble learning for fault diagnosis in refrigeration systems," *Appl. Thermal Eng.*, vol. 164, Jan. 2020, Art. no. 114516.

[37] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[38] D. Mohapatra, B. Subudhi, and R. Daniel, "Real-time sensor fault detection in tokamak using different machine learning algorithms," *Fusion Eng. Des.*, vol. 151, Feb. 2020, Art. no. 111401.

[39] Z. Noshad, N. Javaid, T. Saba, Z. Wadud, M. Saleem, M. Alzahrani, and O. Sheta, "Fault detection in wireless sensor networks through the random forest classifier," *Sensors*, vol. 19, no. 7, p. 1568, 2019.

[40] T. Li, J. Leng, L. Kong, S. Guo, G. Bai, and K. Wang, "DCNR: Deep cube CNN with random forest for hyperspectral image classification," *Multimedia Tools Appl.*, vol. 78, no. 3, pp. 3411–3433, Feb. 2019.

[41] J. M. Malof, L. M. Collins, K. Bradbury, and R. G. Newell, "A deep convolutional neural network and a random forest classifier for solar photovoltaic array detection in aerial imagery," in *Proc. IEEE Int. Conf. Renew. Energy Res. Appl. (ICRERA)*, Nov. 2016, pp. 650–654.

[42] E. Min, J. Long, Q. Liu, J. Cui, and W. Chen, "TR-IDS: Anomaly-based intrusion detection through text-convolutional neural network and random forest," *Secur. Commun. Netw.*, vol. 2018, pp. 1–9, Jul. 2018.

[43] S. Bai, "Growing random forest on deep convolutional neural networks for scene categorization," *Expert Syst. Appl.*, vol. 71, pp. 279–287, Apr. 2017.

[44] Y. Liu, X. Yuan, X. Gong, Z. Xie, F. Fang, and Z. Luo, "Conditional convolution neural network enhanced random forest for facial expression recognition," *Pattern Recognit.*, vol. 84, pp. 251–261, Dec. 2018.

[45] U. Knauer, C. S. von Rekowski, M. Stecklina, T. Krokotsch, T. P. Minh, V. Hauffe, D. Kilias, I. Ehrhardt, H. Sagischewski, S. Chmara, and U. Seiffert, "Tree species classification based on hybrid ensembles of a convolutional neural network (CNN) and random forest classifiers," *Remote Sens.*, vol. 11, no. 23, p. 2788, 2019.

[46] E. Anggiratih and A. E. Putra, "Ship identification on satellite image using convolutional neural network and random forest," *Indonesian J. Comput. Cybern. Syst.*, vol. 13, no. 2, pp. 117–126, 2019.

[47] G. Xu, M. Liu, Z. Jiang, D. Söffker, and W. Shen, "Bearing fault diagnosis method based on deep convolutional neural network and random forest ensemble learning," *Sensors*, vol. 19, no. 5, p. 1088, 2019.

[48] Y. Le Cun, L. D. Jackel, B. Boser, J. S. Denker, H. P. Graf, I. Guyon, D. Henderson, R. E. Howard, and W. Hubbard, "Handwritten digit recognition: Applications of neural network chips and automatic learning," *IEEE Commun. Mag.*, vol. 27, no. 11, pp. 41–46, Nov. 1989.

[49] H. Wu and J. Zhao, "Deep convolutional neural network model based chemical process fault diagnosis," *Comput. Chem. Eng.*, vol. 115, pp. 185–197, Jul. 2018.

[50] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.

[51] J. Günther, P. M. Pilarski, G. Helfrich, H. Shen, and K. Diepold, "First steps towards an intelligent laser welding architecture using deep neural networks and reinforcement learning," *Procedia Technol.*, vol. 15, pp. 474–483, 2014.

[52] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[53] Y. LeCun. (2015). *LeNet-5, Convolutional Neural Networks.* [Online]. Available: http://yann.lecun.com/exdb/lenet

[54] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[55] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, *arXiv:1312.4400*. [Online]. Available: http://arxiv.org/abs/1312.4400

[56] V. T. Do and U.-P. Chong, "Signal model-based fault detection and diagnosis for induction motors using features of vibration signal in two-dimension domain," *Strojniški Vestnik-J. Mech. Eng.*, vol. 57, no. 9, pp. 655–666, Sep. 2011.

[57] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," 2012, *arXiv:1207.0580*. [Online]. Available: http://arxiv.org/abs/1207.0580

[58] H. Qiu Zhang and Y. Yan, "A wavelet-based approach to abrupt fault detection and diagnosis of sensors," *IEEE Trans. Instrum. Meas.*, vol. 50, no. 5, pp. 1389–1396, 2001.

[59] Z. Shen and Q. Wang, "Failure detection, isolation, and recovery of multifunctional self-validating sensor," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 12, pp. 3351–3362, Dec. 2012.
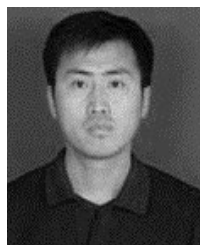
YONGYI SUN received the B.S. degree from Heilongjiang University, in 2004, and the M.S. degree from Northeastern University, in 2009, where he is currently pursuing the Ph.D. degree. He once went to the Computer Science Department, The Hong Kong University, as a Visiting Scholar, in 2019. His main research interests include machine learning, deep learning, and intelligent sensor systems.

HONGQUAN ZHANG received the M.S. and Ph.D. degrees in measurement technology and instruments from the Harbin University of Science and Technology, in 2000 and 2008, respectively. Since 2014, he has been serving as a Doctoral Tutor with both Heilongjiang University and Harbin Engineering University. His main research interests include application of sensitive materials, functional materials, micro-processing technology research, new sensor design, and manufacturing methods.

TINGTING ZHAO received the B.S. degree in automation engineer from Harbin Engineering University, in 2019, where she is currently pursuing the M.S. degree in instrument science and technology. Her research interests include sensor networks, data fusion, and fault detection and diagnosis.

ZHIHUI ZOU received the B.S. degree in measurement and control technology and instrumentation from Harbin Engineering University, in 2019, where he is currently pursuing the M.S. degree in control science and engineering. His research interests include sensor design, intelligent control, and data fusion.

BIN SHEN received the B.S. degree in computer science and technology and the M.S. degree in safety technology and engineering from the Heilongjiang University of Science and Technology, in 2006 and 2009, respectively. He is currently pursuing the Ph.D. degree with Harbin Engineering University. He was a Teacher with the Heilongjiang University of Science and Technology. His research interests include gas sensor, wireless sensor network, new sensor design, and manufacturing methods.

LIXIN YANG received the B.S. degree in mathematics and applied mathematics and the M.S. degree in control theory and control engineering from Heilongjiang University, in 2005 and 2008, respectively, and the Ph.D. degree in systems engineering from Harbin Engineering University, in 2012. Her research interests include control systems, pattern recognition, and the signal processing of gas sensor arrays.

• • •