FD Technology for HSs based on Deep Convolutional Generative Adversarial Networks

Jun Wang
College of Electronic and Information Engineering, Tongji University, China
Wangjun13WJ@outlook.com

Yuanxi Wang
College of Electronic and Information Engineering, Tongji University, China
wang_yuanxi@outlook.com

Abstract: *To ensure the reliability of sensors, it is very important to study Fault Diagnosis (FD) methods for sensors. This study puts forward a Convolutional Neural Network (CNN)-FD model with randomly discarding network units. This model reduces the huge computational burden caused by excessive parameters in the CNN through silent neural nodes to achieve efficient automatic extraction and analysis of fault features. Considering the signal data imbalance caused by small sample failures, this study used Generative Adversarial Networks (GANs) to achieve intelligent expansion of samples. The performance test results showed that the proposed model achieved the highest classification accuracy in binary classification tasks, with a size of 93.5%, which is 5.5% and 3.5% higher than the Densenet model and ResNet model, respectively. In multi-classification tasks, the model still realized the best classification accuracy, with a size of 89.9%, which is 8.8% higher than the Densenet model. The experimental results of fault detection denoted that the proposed model arrived the highest recognition accuracy in aging failure and sensitivity failure, with 96.8% and 92.6% respectively, while the recognition accuracy of shallow neural networks and deep confidence networks was lower than 90%. This indicated that the proposed algorithm can effectively perform feature extraction and fault pattern recognition.*

Keywords: *CNN, GAN, hydrogen gas, sensors, fault, pattern recognition.*

Received December 13, 2023; accepted February 11, 2024
https://doi.org/10.34028/iajit/21/2/12

1. Introduction

Energy is the foundation of a country's development, and non-renewable resources such as oil, natural gas, and coal are important factors driving the modernization of the global economy [14]. However, if the oil and natural gas resources are exhausted due to excessive consumption of oil and natural gas, the global energy crisis will be triggered. Under the theme of sustainable development, it is necessary to research and use cleaner energy sources. Hydrogen is a gas composed of the lightest elements and has high dispersibility, so its propagation speed in the atmosphere is much faster than gasoline [32]. As soon as hydrogen gas leaks, it will quickly spread outward. Its highly dispersed nature makes it more difficult to control compared to other gases. Moreover, the combustible range of hydrogen is wider, so scientifically and efficiently measuring it to reduce or avoid problems such as explosion is currently an urgent problem to be solved [1]. Hydrogen Sensors (HSs) are a primary means of effectively detecting hydrogen content and quickly capturing and detecting hydrogen in the event of hydrogen leakage. When the hydrogen content exceeds the normal range, it can also use a secondary instrument to alarm. Semiconductor HSs have been used as a gas sensor for the determination of hydrogen content due to their advantages such as low price, high sensitivity, and easy operation. However, in long-term use, the detection characteristics of semiconductor HS components are not only related to hydrogen content, but also closely related to factors such as temperature, humidity, and pressure. In engineering practice, due to environmental temperature and humidity, vibration and shock, the sensitivity of HSs is prone to degradation and failure [30]. Under real working conditions, changes in gas concentration may occur due to the influence of air humidity or the falling of gaseous substances. Not only that, due to external vibrations and other reasons, it is easy to break the heating and the lead wires, thereby causing the output of the sensor to malfunction. These factors will ultimately cause parameter deviations of sensors, reduce system efficiency, and even cause sensor failure, resulting in incorrect measurements and judgments, leading to significant disasters [31]. Therefore, Fault Diagnosis (FD) of HSs is very necessary. In view of this, this study proposed an optimized Convolutional Neural Network (CNN) model for feature extraction and learning of sensor faults to achieve accurate FD. In addition, considering the imbalanced fault data, this study also combined the Generative Adversarial Networks (GANs) to generate high similarity fault samples to improve the recognition effect of multi category faults.

This study is mainly composed of four parts for discussion. The first section is a literature review on CNNs, GANs, and sensor faults. The second section is a fault feature recognition model constructed using an

improved CNN model and a GAN used for fault small sample expansion. The third section is a performance and application analysis of deep convolutional GANs. The fourth part is a conclusive summary of the FD model proposed by the research institute. The reason for using CNN improved based on dropout strategy is that it shares weights, reduces the number of links in the network, and reduces the risk of overfitting. The reason for using GANs based on duality generation is that it helps to enhance the sample dataset and generate artificially synthesized data samples.

2. Related Works

Many scholars have conducted extensive research on CNNs or GANs, their research results are shown in Table 1.

Table 1. Comparison of previous research achievements on GANs.

Number	Literature title	Research findings	Research deficiency
[20]	Retinal vessel segmentation using Multi-Scale Residual convolutional neural Network (MSR-Net) combined with GANs	Kar *et al*. [20] propose a method for accurately detecting retinal blood vessels from fundus images using GAN's multiple loss function	Small sample size for testing
[27]	CNNs as a model of the visual system: past, present, and future	Lindsay [27] elaborates on the significance of cellular neural networks for visual tasks and discusses new opportunities for cellular neural networks in visual research beyond basic object recognition	For quantitative scientific validation
[7]	Multi-scale GAN for image super-resolution	Daihong *et al*. [7] propose a multi-scale GAN to address the difficulty of reconstructing high-frequency information and details in low resolution images	The algorithm has low computational efficiency
[21]	Generating three-dimensional structures from a two-dimensional slice with GAN-based dimensionality expansion	Kensh and Cooper [21] design a GAN architecture called Slice GAN, which can synthesize high fidelity 3D datasets using a single representative 2D image	Slow processing speed
[24]	Lost data reconstruction for structural health monitoring using deep convolutional GANs	Lei *et al*. [24] propose a deep convolutional GAN	There are fewer types of signals that can be processed
[8]	Sensor-fault detection, isolation and accommodation for digital twins via modular data-driven architecture	Darvishi *et al*. [8] propose a sensor validation architecture based on general machine learning to detect anomalies in sensor measurements and identify erroneous anomalies	Small sample size for testing
[10]	Sensor fault detection and isolation via networked estimation: Full-rank dynamical systems	Doostmohammadian and Meskin [10] propose a sensor replacement scheme based on graph theory	/
[18]	Sensor fault detection and isolation using a SVM for vehicle suspension systems	Jeong *et al*. [18] propose a residual generation method based on fault isolation observer	Operators need to have a certain industry basic knowledge reserve
[38]	Intermittent sensor fault detection for stochastic LTV systems with parameter uncertainty and limited resolution	Zhang *et al*. [38] designed a sensor fault detection model for stochastic Linear Time-Varying (LTV) systems	The model is black box and has poor interpretability
[33]	Event-triggered fuzzy filtering for networked systems with application to sensor fault detection	Tan *et al*. [33] propose an event triggering scheme with two triggering matrices,	Narrow application range
[36]	Dropout technique for image classification based on extreme learning machine	Wen *et al*. [36] propose a new CNN learning rate scheduler based on reinforcement learning to enhance the potential of fault classification for classifying fault data diversity.	Fewer training samples
[15]	Research and application of deep learning in image recognition	Guo *et al*. [15] propose a deep learning based FD method for small current grounded distribution systems	Slow recognition speed
[19]	A new reinforcement learning based learning rate scheduler for CNN in fault classification	Jiang *et al*. [19] propose a CNN based azimuth FD method to eliminate noise interference and consider possible connections between signal frames	Insufficient test data volume
[23]	Deep-learning-based fault classification using Hilbert-huang transform and CNN in power distribution systems	Kiranyaz *et al*. [23] discuss and compares the performance and advantages and disadvantages of 1D CNN and 2D CNN in classification tasks	Lack of dataset based comparative testing
[2]	Distributed Reduced CNNs	Alajanbi *et al*. [2] design a distributed minimal CNN algorithm	/

Kar *et al*. [20] proposed an accurate retinal blood vessel detection method from fundus images by using multiple loss function of the GAN. The GAN structure included a generator and a discriminatoras segmentation and classification networks, respectively. The outcomes proved that the accuracy of the method on multiple datasets exceeded 90%. Lindsay [27] elaborated on the significance of CNNs for visual tasks and discussed new opportunities for cellular neural networks in visual research beyond basic object recognition. Daihong *et al*. [7] proposed a multi-scale GAN to address the difficulty in reconstructing high-frequency information and details in low-resolution images. The multi-scale pyramid module inside the generator could extract features containing high-frequency information, and then used bicubic interpolation results to reconstruct high-resolution images. The outcomes demonstrated that the algorithm performed better on the two indicators of super-resolution tasks, Peak Signal-to-Noise Ratio PSNR and Structural Similarity Metric (SSIM). Kench and Cooper [21] introduced a GAN architecture called slice GAN, which could synthesize high-fidelity 3D datasets using a single representative 2D image. This was particularly important for the task of generating material micro structures, as cross-sectional micro graphs could contain sufficient information to statistically reconstruct 3D samples. To reconstruct the lost data in the structural health monitoring, Lei *et al*. [24] proposed a deep convolution generation countermeasure network, which included a generator with a codec structure and a countermeasure discriminator. The results showed that the final reconstructed signal matched well with the real signal in both the time and frequency domains.

Many scholars have adopted various research methods in sensor FD. Darvish *et al*. [8] proposed a universal machine learning based sensor validation architecture to detect anomalies in measurements from sensors and identify erroneous anomalies. The comprehensive statistical analysis results on three different real datasets showed that this method performed well in identifying soft and hard comprehensive faults. Doostmohammadian and Meskin [10] have considered the simultaneous sensor fault detection, isolation, and networked estimation in linear full-rank dynamic systems, and raised a sensor replacement scheme based on graph theory to recover potential network observability losses caused by the removal of faulty sensors. The study examined the fault detection and isolation scheme on an illustrative academic example to validate the outcomes and compare them with relevant literature. Jeong *et al*. [18] proposed a method based on fault isolation observers to generate residuals and use Support Vector Machines (SVMs) for estimation. The results showed that the FD algorithm was desired to lessen the workload required in the design process, and could also detect a small number of sensor faults. Zhang *et al*. [38] considered the detection of intermittent sensor faults in stochastic linear time-varying systems with parameter uncertainty and limited resolution. Through recommending a soft sensor model, a state estimator was designed, where the upper bound of the estimation error covariance was minimized at each time step. Finally, this method's effectiveness was validated through two simulation experiments. Tan *et al*. [33] proposed an event triggering scheme with two triggering matrices to study the filtering of event triggered fault detection in nonlinear networked control systems. Finally, the new fault detection filtering technology's effectiveness and superiority were demonstrated through examples.

In the application of CNN in the direction of fault classification, Wen *et al*. [36] proposed a new CNN learning rate scheduler based on reinforcement learning to improve the potential of fault classification for the diversity classification of fault data. The results indicated that the algorithm performance was better than traditional machine learning methods. Guo *et al*. [15] proposed a FD method for small current grounding distribution system based on deep learning, and used image similarity recognition method based on CNN to classify faults. The findings expressed that this method hadhigh accuracy and adaptability in FD of distribution systems. To eliminate noise interference and consider possible connections between signal frames, Jiang *et al*. [19] proposed a bearing FD method based on CNN. A filtering method based on spectral kurtosis was proposed to suppress noise by utilizing the sensitivity of spectral kurtosis to pulses. Kiranyaz *et al*. [23] summarized that 1D CNN has recently become the latest technology of key signal processing applications, such as patient specific ElectroCardioGram (ECG) classification, structural health, power electronic circuit anomaly and motor fault detection. The difference between it and 2D CNN was discussed.

In summary, scholars mainly use CNNs or GANs to improve image quality or reconstruct images. In sensor FD, they mainly use residual networks or SVMs for analysis, and rarely combine CNN and GAN for gas sensor fault detection. Alajanbi *et al*. [2] found that kernel based CNN algorithms can solve nonlinear supervised tasks, but this algorithm is time-consuming and memory intensive when dealing with large-sized kernel matrices. Therefore, they designed a distributed minimalist CNN algorithm that uses a distributed minimalist kernel to store data from multiple locations, and uses distributed training techniques based on alternating direction multipliers for model training. The test results show that the training and computation time of the improved algorithm is much lower than that of the original algorithm.

Existing research mainly focuses on using CNN or GAN models for fault signal recognition or estimation, but there are few studies that use them to optimize fault data to achieve a balanced effect.

For this, this study puts forward a deep convolutional GAN to diagnose HS faults and improve its diagnostic accuracy. The novel design of CNN technology is optimized using the dropout strategy. The dual GAN proposed by the research institute has noise generation and denoising networks, which are continuously enhanced through mutual dual regularization and optimized using the Wassertein distance function, which is different from traditional GAN networks.

3. Construction of Sensor FD Model Based on Deep Convolutional Kernel GAN

Due to different working conditions, the characteristics and quantity of fault signal data obtained by HSs also vary [22, 26, 35]. Current pattern recognition methods such as SVMs or extreme learning machines require manual feature extraction, resulting in uncertainty in the quality of extracted features [28]. To extract features without relying on expert experience, this study proposes an improved CNN algorithm to implement an end-to-end HSFD system. In addition, there is an imbalance in the sample size of different fault data [11, 13]. To ensure that the network can train well for faults with small sample sizes, this study introduces GANs to optimize the small sample problem.

3.1. Improved CNN FD Model Based on Random Dropping of Network Units

The neurons of CNNs can be optimized through training and can also perform single input independent operations [5]. The neurons distributed in different layers of CNNs contain three spatial dimensions: height, width, and depth. CNNs typically perform data

processing based on four principles: internal links, weight sharing, pooling, and multi-layer [4, 25]. The CNN structure is sequentially divided into input, convolutional, pooling, and fully connected layers [37]. By continuously training CNNs, the optimal model structure can be obtained. Convolutional layers have a crucial influence on the effectiveness of CNNs. In sensor FD, they can extract different input fault signal features. The feature map of each fault signal is composed of a rectangular neural network, and the neural network of a single fault signal has a common weight called a convolutional kernel [29]. The convolutional kernel is usually initialized as a random matrix, and the convolutional layer is shown in Figure 1.

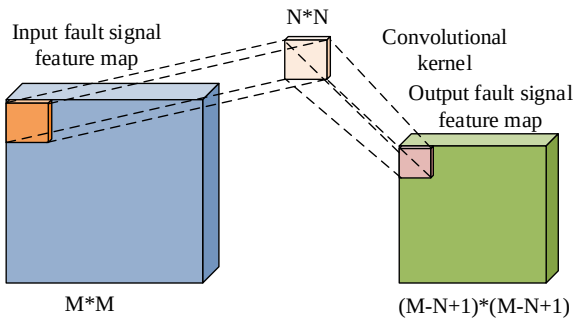


Figure 1. Convolutional layer.

The convolution layer contained in the CNN has several filters, which is a predefined hyper parameter [3, 34]. The number of filters within a layer represents the depth dimension of the output of the activated feature map, which is created by the convolutional layer as input to the next layer [17]. Each filter has a set width and height, corresponding to the local receiving field of a single unit within the layer. The filter acting on the input data produces an output of a convolutional layer, known as feature mapping. In the training phase of CNN, the weight values in the filter can be learned [39]. The output dimension of the convolutional layer has a depth component, and if each segment of the output is segmented, a feature map of a two-dimensional plane will be obtained. A filter used on a single two-dimensional plane contains a weight that is shared among all filters used on the same plane. This maintains the same feature detector in another part of the input data as in another part of the input data. Therefore, the direct advantage of using common weights is that it reduces the number of links in the network and the risk of overfitting [6]. Z_k means the eigenvalues of the K-th feature map of a nonlinear transformation, which are calculated by Equation (1).

$$Z_k = W_k \otimes \tilde{x} + b_k \tag{1}$$

In Equation (1), $\tilde{x}$ is the two-dimensional grayscale image of the input sensor fault signal; W_k expresses the convolution kernel of the sensor fault signal feature map; b_k denotes the offset; $\otimes$ refers to the convolution of the two-dimensional grayscale image; H indicates a two-dimensional convolutional kernel, with input I. The convolution calculation is represented by Equation (2).

$$S(i,j) = (I \times H)(i,j) = \sum_m \sum_n I(m,n)H(i-m, j-n) \tag{2}$$

In Equation (2), $S(i,j)$ denotes the result of convolution calculation. The pooling layer, also known as subsampling, is used for down sampling the characteristic map of the input sensor. Pooling is an important concept in CNN, which is actually a form of down sampling. There are various forms of nonlinear pooling functions, among which maximum pooling is the most common. It divides the input image into several rectangular regions and outputs the maximum value for each sub region. This mechanism is effective, because after discovering a feature, its precise position is far less important than its relative position with other features. The pooling layer continuously reduces the spatial size of the data, resulting in a decrease in the number of parameters and computational complexity, which to some extent also controls overfitting. Due to the absence of parameters in the pooling layer, this method can cut down the dimensionality of the feature expression of the fault signal and computational time and net parameters. This study applies the maximum pool function to calculate the maximum value in a local neighborhood rectangle. In the pooling layer, assuming the characteristic map of a sensor fault as input, Equation (3) is used to calculate the output of the level I characteristic map.

$$x_j^l = f\left(\beta_j^l down\left(x_j^{l-1}\right) + b_j^l\right), j = 1, \ldots, N \tag{3}$$

In Equation (3), b_j^l expresses additive deviation; β_j^l means multiplicative deviation; x_j^{l-1} infers to input diagram of j layer; x_j^l stands for output diagram of j layer; f indicates activation function; *down* denotes sub sampling function. Figure 2 is a schematic diagram of the fully connected layer.

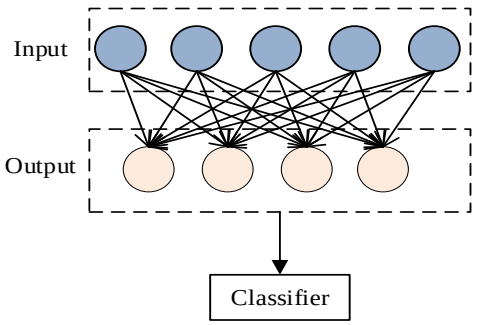


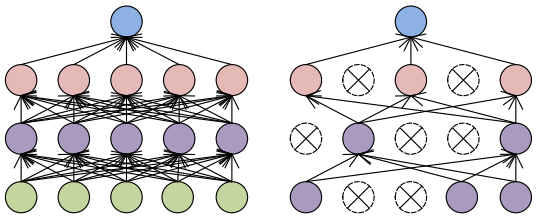
Figure 2. Schematic diagram of fully connected layer.

The convolutional and pooling layers are used to extract the characteristics of sensors, while the fully connected layer is used to convert the feature information of 2D grayscale images into 1D feature vectors. The Softmax classifier makes final judgments

and identifies fault patterns based on the characteristics of faults, and outputs FD results. Equation (4) is the expression for the output result of the classifier.

$$y(a_i) = \frac{exp(a_i)}{\sum exp(a_i)} \tag{4}$$

In Equation (4), *y(ai)* represents the output result of the classifier; *a* is the output value of the fault signal feature vector in the fully connected layer. Because of the massive parameters in the CNN network, it will increase computational complexity and slow down computational efficiency. To minimize computational complexity as much as possible, this study adopts a network unit random discarding method, as shown in Figure 3.



a) Traditional neural network. b) Dropout neural network.

Figure 3. Neural network model diagram with random discarding of network units.

Random discarding is a method of implementing different combinations of nodes in a neural network by actively and randomly silencing some nodes (while still retaining their weight values) [12]. The key concept of this method is to randomly remove units from the neural network to avoid excessive adaptation between units [16]. In addition, it is an effective method to approximate the exponential combination of various neural networks [9]. Removing a unit from the network means temporarily removing any input and output links of that unit from the network. Equation (5) is the expression for calculating the feedforward operation when network units are randomly discarded.

$$\begin{cases} r_j^{(l)} \leftarrow Bernoulli(p) \\ \tilde{y}^{(l)} = r^{(l)} \times y^{(l)} \\ z_i^{(l+1)} = w_i^{(l+1)} y + b_i^{(l+1)} \\ y_i^{(l+1)} = f\left(z_i^{l+1}\right) \end{cases} \tag{5}$$

In Equation (5), l denotes the hidden layer number; $z_i^{(l+1)}$ is the input vector; $y^{(l)}$ represents the output vector; $w_i^{(l+1)}$ stands for the weight value; $b_i^{(l+1)}$ means offset; f stands for activation function; × infers to the multiplication of vectors by their corresponding elements; $r_j^{(l)}$ denotes a random variable and obeys the Bernoulli distribution; $\tilde{y}^{(l)}$ indicates the product of the random variable and the output $y^{(l)}$ of the layer, and is used as the input value for the next layer. z means the category value for sensor fault mode diagnosis.

3.2. Construction of Small Sample Fault Equilibrium Model Based on Dual GAN

The feature extraction algorithm based on traditional deep learning relies on a large dataset for model training. And the dataset it uses belongs to paired datasets with and without noise, which greatly increases the learning time of the model and reduces the efficiency of image extraction. Automatic generation of datasets is one of the functions that GAN can achieve. For fault feature images, the generating noise samples by fitting the initial image or the mapping relationship between fitting noise and non-noise samples is the key to dataset generation. However, the disadvantage of these two methods is that the former has too many hyper parameters, which requires a lot of manual adjustment and processing, while the latter has unstable learning due to too high learning dimensions. Therefore, the study chooses a dual GAN to establish a joint distribution relationship between the noiseless and noisy images, which avoids the need for a large dataset to construct the mapping relationship between the two images. Figure 4 is a schematic diagram of the network structure composition.

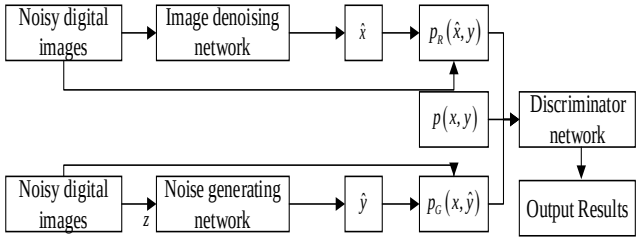


Figure 4. Composition diagram of dual GAN structure.

From Figure 4, the network mainly consists of three functional sub-networks for noise generation, image denoising, and sample authenticity discrimination. The generation of noise and image denoising belongs to the reciprocal process, and this reciprocal relationship is the basis for establishing joint distributions in the network. Research has recorded them as $p_R(x, y)$ and $p_G(x, y)$, respectively. The enhanced noise samples automatically generated by the network will train the image denoising network in the future, which is beneficial for improving the denoising level of the network. Considering that UNet networks have the advantage of occupying fewer computing resources and having higher learning speed, this network module is used to construct noise generation networks and image denoising networks. The discriminator is composed of a single-layer fully connected network and a multi-layer CNN. It assumes that x and y represent the initial fault feature image without and with noise, respectively, then digital image data will be obtained for (x, y). Equation (6) is the transformation expression of the input data pair.

$$\begin{cases}\hat{x}=R(y)\\ \hat{y}=G(x,z)\end{cases} \tag{6}$$

In Equation (6), $\hat{x}$ and $\hat{y}$ express the enhanced sample data output through the noise generating network *G* and the denoising network *R*. This enhancement of sample data comes from the dual regularization in the dual GAN. Since $\hat{x}$ needs to be given noise information, the variable parameter *z* that obeys normal distribution and belongs to isotropy is introduced to simulate the noise brought by the hardware itself in the image generation. x, y, $\hat{x}$, and $\hat{y}$ can form data pairs $(x,\hat{y})$ and $(\hat{x},y)$, and then $p_R(\hat{x},y)$ and $p_G(x,\hat{y})$ are obtained using Bayesian thinking. The mathematical expressions for networks *R* and *G* are in Equation (7).

$$\begin{cases}p_R(x,y)=p_R(x|y)\,p(y)\\ p_G(x,y)=\int_z p_G(y|x,z)\,p(x)\,p(z)\,dz\approx\frac{1}{L}\sum_i^L p_G(y|x,z_i)p(x)\end{cases} \tag{7}$$

In Equation (7), $p_R(x|y)$ and $p_G(y|x,z)$ denote the conditional probability distribution, and *p(x)*, *p(y)*, and *p(z)* represent the general probability distribution. The probability based joint distribution $p_R(x, y)$ and $p_G(x, y)$ can gradually approximate the true joint distribution relationship *p(x, y)* of the image after adversarial learning. Figure 5 is a schematic diagram of the UNet network framework.

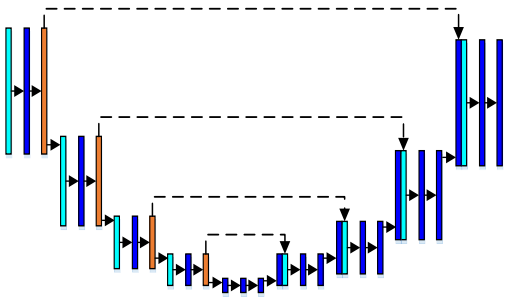


Figure 5. Schematic diagram of UNet network framework.

From Figure 5, the UNet network mainly performs image scaling and restoration processing. In the image scaling operation, the linear convolution layer is utilized to control the size of the image feature map. Image restoration processing utilizes image transposition to transform the feature map into its original size. In the learning of image features, the network adopts a residual learning method, as shown in Equation (8).

$$\begin{cases}R(y)=y-U(y)\\ G(x.z)=x+U([x,z])\end{cases} \tag{8}$$

In Equation (8), *U(·)* denotes the forward propagation algorithm. The discriminator network is selected to replace the general loss function in the dual GAN. Figure 6 is a schematic diagram of the discriminator network structure.

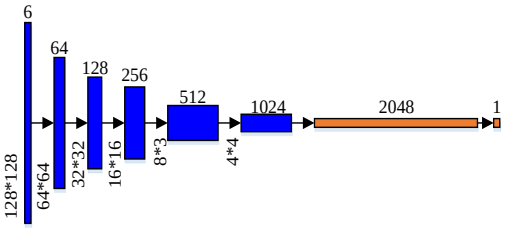


Figure 6. Schematic diagram of discriminator network structure.

From Figure 6, the discriminator network has a linear convolutional layer of 5 and a fully connected layer of 1. The function of linear convolutional layers is to lessen the size of feature maps, while the function of fully connected layers is to blend image features. The function of the entire discriminator is to judge between real and false images, so that the adversarial network is in the correct training direction.

3.3. GAN Model Optimized Based on Wassertein Distance Function

When identifying fault features, it mainly includes three steps: extracting real noise, calculating noise residuals, and feature estimation. Several digital images are obtained through the device, and then denoising operations are performed on them using a dual GAN to obtain the denoised digital images. Afterwards, the noise residual is calculated using Equation (9).

$$W^i=I^i-C^i \tag{9}$$

In Equation (9), I^i refers to the original digital image, and C^i denotes the denoised digital image. The maximum likelihood estimation value $\hat{K}$ of fault features can be obtained using residual values, and its calculation is shown in Equation (10).

$$\hat{K}=\frac{\sum_{i=1}^{M}W_iI_i}{\sum_{i=1}^{M}\left(I^i\right)^2} \tag{10}$$

Although the dual GAN has advantages in learning image sample features, the gradient vanishing situation that exists in the GAN itself has not disappeared. Due to the interruption of network training caused by the disappearance of gradients, the model loses its image recognition function. Therefore, it is necessary to optimize the above model. The study adopts Wassertein distance as the similarity training objective to improve the gradient vanishing problem. Equation (11) is a mathematical expression for calculating the Wassertein distance.

$$\min_{R,G}\max_{D}\ell_{gan}(R,G,D)=E_{(x,y)}\left[D(x,y)\right]-\alpha E_{(\hat{x},y)}\left[D(\hat{x},y)\right]-(1-\alpha)E(x,\hat{y})\left[D(x,\hat{y})\right] \tag{11}$$

In Equation (11), *D* means the discriminator, and α is the hyper parameter used to allocate the weights of the

noise generation and the noise removal networks. The purpose of calculating the distribution distance is to understand the distinguish between the actual output image and the expected output image. The similarity measurement parameters originally used to GANs are Kullback Leibler (KL) divergence or Jensen Shannon (JS) divergence, and their mathematical calculation expressions are shown in Equation (12).

$$\begin{cases} KL\left(P\|Q\right)=\sum\limits_{x\in X}P\left(x\right)\log\dfrac{P\left(x\right)}{Q\left(x\right)} \\ JS\left(P_1\|P_2\right)=\dfrac{1}{2}KL\left(P_1\left\|\dfrac{P_1+P_2}{2}\right.\right)+\dfrac{1}{2}KL\left(P_2\left\|\dfrac{P_1+P_2}{2}\right.\right) \end{cases} \tag{12}$$

In Equation (12), *P(x)* and *Q(x)* indicate the probability distributions of real and theoretical existence events, respectively. Although *KL* or *JS* have the function of characterizing the similarity between images, both divergences use the log function for similarity calculation. This causes the similarity value to suddenly change to a constant log2 as the discriminator is trained, leading to the phenomenon of gradient disappearance. The Wassertein distance function, on the other hand, has fewer fluctuations in its value and tends to be more stable during training due to the absence of mutation points. In addition to replacing the similarity measurement equation to optimize the original network, the study also considers introducing the loss function to speed up the network training. Equation (13) is the expression of *loss* function.

$$LOSS=\left\|\varsigma F\left(\hat{y}-x\right)-\varsigma F\left(y-x\right)\right\|_1 \tag{13}$$

In Equation (13), $\varsigma F(\cdot)$ indicates a Gaussian filter. Equation (13) is based on $loss\left\|\hat{x}-x\right\|_1$. The reason why $\left\|\hat{x}-x\right\|_1$ is not directly used as the loss function is that the noise is a strong random signal. Figure 7 is a schematic diagram of the image update.

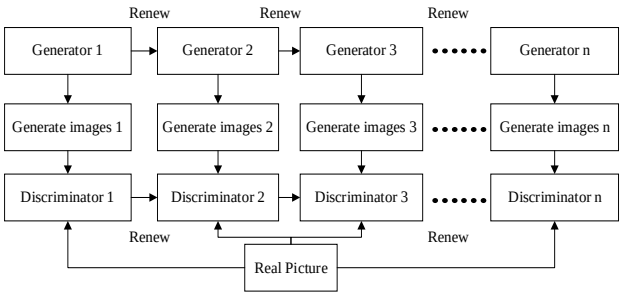


Figure 7. Image update process.

From Figure 7, the image output by the generator will be fed into the discriminator for true or false judgment. The goal of the generator is to obtain an image that is closer to the expected output, while the goal of the discriminator is to identify whether the image is false. As a result, a confrontational relationship is formed between the two. When the image output by the generator is sufficiently realistic to deceive the discriminator, both will no longer be updated, and an ideal network model will be obtained.

4. Performance Evaluation Experiment and Utility Analysis of Deep Convolutional GAN Algorithm

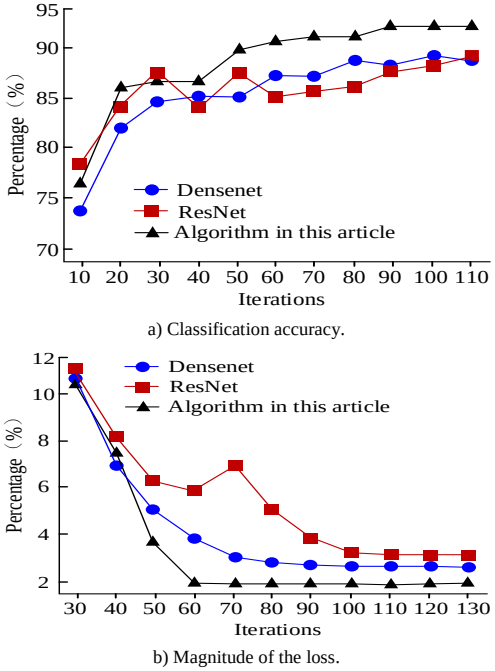
To verify the superiority of the deep convolutional GAN algorithm proposed by the research institute and its effectiveness in sensor FD, this study set up binary and multi-classification tasks to evaluate the algorithm performance. And five types of sensor fault signals are selected as experimental data to evaluate the algorithm's application effectiveness.

4.1. Analysis of the Performance Results of Deep Convolutional GAN Algorithms

The CIFAR-10 dataset was collected by professors from the University of Toronto in Canada and is a common database for tasks in the fields of machine learning and image recognition. The database contains images of airplanes, cars, cats, birds, dogs, deer, frogs, horses, boats, and trucks, all of which are 32 images×32 size color image. Due to the diverse color images included in the CIFAR-10 dataset and the image specifications meeting the requirements of the experimental environment, the performance evaluation experimental dataset was selected as CIFAR-10, this dataset was used for performance comparison analysis between the designed algorithm and the comparative algorithm in this study. This dataset contained 6*104 color images, with image types covering vehicles, animals, etc. The proportion of training and testing samples included was 5:1. The grid structure used in the experiment was uniform. The convolution kernel of the grid structure was set to 4*4; the pooling layer used the maximum pooling function; the step size was set to 2; the sliding window was set to 2*2; the random abandonment probability of neurons was set to 0.5; the activation function was set to ReLU; and the learning rate was set to 0.005. Figure 8 shows the classification accuracy and loss value curves of different models in the binary classification task. In this study, the consistency between the predicted image type output by the model and the actual label will be used as the criterion for judging the correctness of classification.

From Figure 8-a), the classification accuracy curve corresponding to the model proposed in the study had the fastest climbing speed,and was in a stable growth state, with little fluctuation. The classification accuracy has already exceeded 90% when the iteration times was 50, and the convergence value was 93.5% when the amount of iterations was 110. The Densenet model corresponded to the slowest growth rate of the classification accuracy curve, and that of the entire curve was below 90%. When the amount of iterations

was 110, the classification accuracy was 88%, which was 5.5% lower than the former. The classification accuracy curve of the ResNet model fluctuated significantly in the early stage, but remained stable in the later stage. The final classification accuracy value was 90%, which was 3.5% lower than the model proposed in the study. From Figure 8-b), the proposed model achieved the minimum loss value of 2%. The loss value curve had the fastest descent speed, after 60 iterations, it decreased from 10.3% of the initial loss value to 2.1%, and began to enter the convergence stage. The convergence curve was stable without oscillation, and the convergence performance was good. The initial loss value curve of the Densenet model was 11.2%. When the number of iterations was 100, the curve entered the convergence stage, and the loss value was 2.8%. Afterwards, the curve achieved a final loss value of 2.8% in stationary convergence. The ResNet model had the worst convergence performance, with significant fluctuations in the curve during the decline, and the resulting loss value was relatively large, with a magnitude of 3.2%. By comparison, the proposed model reduced the loss value by 0.8% and 1.2% compared to the latter two models. This indicated that the proposed model had higher classification accuracy and faster convergence speed. Figure 9 shows the classification accuracy and loss value curves of different models in multi-classification tasks.

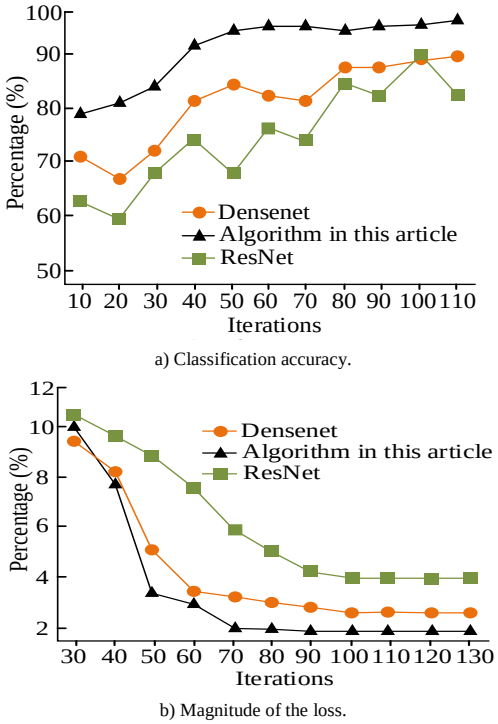


a) Classification accuracy.

b) Magnitude of the loss.

Figure 8. Classification accuracy and loss value curves of different models in binary classification tasks.

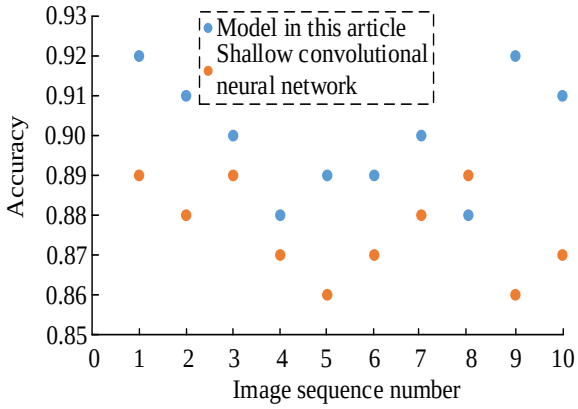
From Figure 9-a), for multi-classification tasks, the classification accuracy curve corresponding to the Densenet model showed a brief downward trend during the iteration. For example, when the iteration interval was [10, 20], the classification accuracy decreased from 70.9% to 66.8%. Then at the iteration interval [50, 70], the classification accuracy decreased from 84.6% to 81.3%. When iterations were 110, the final classification accuracy obtained was 89.9%. For the classification accuracy curve of the ResNet model, because of the drastic fluctuations in the iteration, the convergence value was ultimately not achieved. The classification accuracy curve of the model proposed by the research institute had almost no fluctuations, and it began to converge at the iteration number of 50. The convergence value obtained at the completion of the iteration was 98.7%, which was 8.8% higher than the Densenet model. From Figure 9-b), the loss value curve corresponding to the proposed model showed a rapid decline and then tended to a stable convergence state. The initial loss value of the loss value curve was 10%. When the iteration times was 50, the loss value quickly decreased to 3.7%, and then began to converge at 70. Finally, the minimum loss value of 1.9% was achieved at 130 iterations. The convergence speed of the Densenet model was slower than that of the proposed model. When the iterations were 60, the loss value decreased to 3.8%. Afterwards, the loss value slowly decreased to 2.6%, and finally achieved a convergence loss value of 2.5% when the iteration times was 130. The ResNet model had the slowest descent speed and only converged when the number of iterations was 100, resulting in a final convergence value of 3.9%. By comparison, the proposed model has reduced the loss value by 0.6% and 2.0% compared to the latter two models. This indicated that the proposed algorithm outperformed other algorithms in multi -classification tasks. Figure 10 shows the accuracy comparison scatter plot.



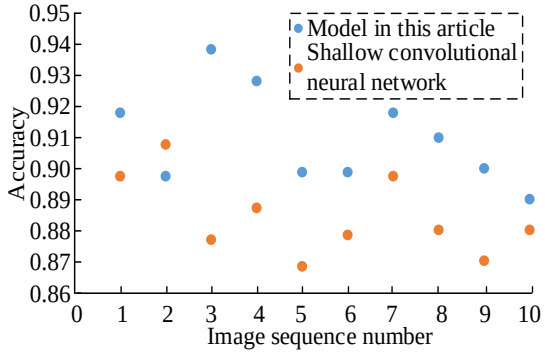
a) Classification accuracy.

b) Magnitude of the loss.

Figure 9. Classification accuracy and loss curve of different models in multi classification tasks.



a) Time classification tasks.

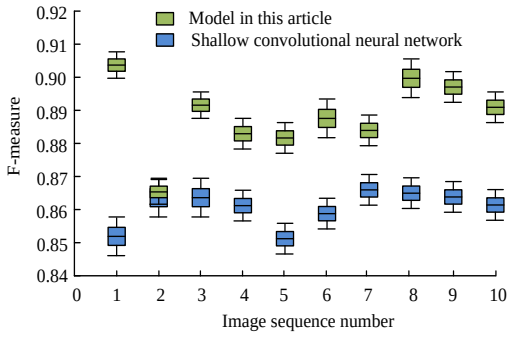


b) Multiple classification tasks.

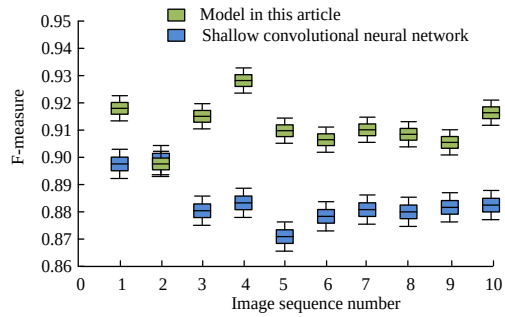
Figure 10. Accuracy comparison scatter plot.

From Figure 10-a), in the accuracy calculation results of the binary classification task, the image recognition accuracy of the proposed model was only lower than that of the shallow CNN model on image 8, while on the other 9 images, the accuracy was higher than the latter. Among them, the accuracy of images 1, 2, 3, 7, 8, and 9 was higher than 0.90, with sizes of 0.92, 0.91, 0.90, 0.90, 0.92, and 0.91, respectively, which was 0.03, 0.03, 0.01, 0.02, 0.06, and 0.05 higher than that of shallow CNNs. From Figure 10-b), in the accuracy calculation results of multi-classification tasks, the model image recognition accuracy used in the study was only lower than that of shallow CNNs on indicator 2, and higher than that of shallow CNNs on the other 9 indicators. The proposed model was higher than 0.90 in images 1, 3, 7, and 8, with the highest value obtained in image 3 being 0.94. The shallow CNN only had an accuracy higher than 0.9 on image 2, with a value of 0.91, which was 0.03 lower than the former. Figure 11 shows the F-metric box plot.

From Figure 11, the proposed model achieved higher indicator values in both binary and multi-classification tasks for the F indicator. For example, for binary classification tasks, the model used in the study had an F indicator higher than 0.86 on all images, with a minimum value of 0.867 obtained in image 2 and a maximum value of 0.91 obtained in image 7. The shallow CNN was not higher than 0.87 on the entire image, indicating that the recognition performance established by the proposed model was better. Figure 12 shows the comparison results of ablation experiments.



a) Time classification tasks.



b) Multiple classification tasks.

Figure 11. F-metric comparison box plot.

From Figure 12, the F-value curve optimized using the dropout strategy for CNN was higher than the unoptimized CNN curve. At the initial value, dropout CNN obtained an F-value of 0.874, while traditional CNN obtained an F-value of 0.862. In terms of F-value convergence, dropout CNN improved by 0.012. This indicated that the proposed model has improved classification accuracy due to the introduction of a random discard strategy.

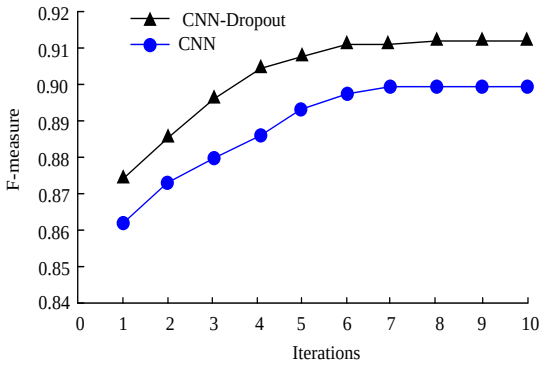
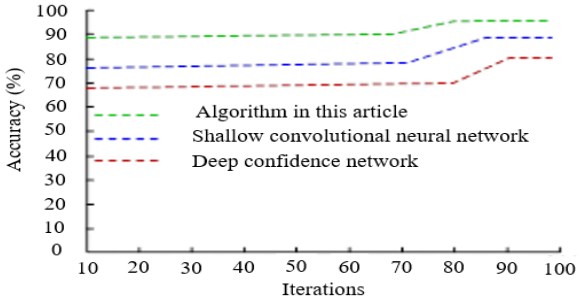


Figure 12. F-metric comparison results.

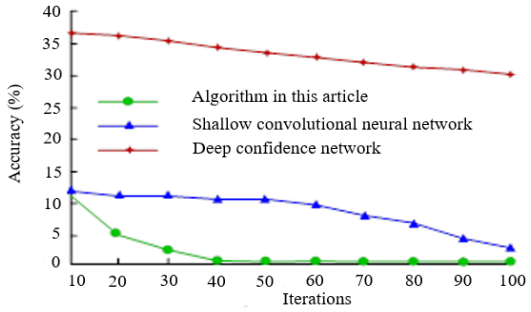
4.2. Analysis of FD Effectiveness of Deep Convolutional GAN Algorithm

Five types of fault sensors were used in the experiment to correspond to aging failure, sensitivity failure, false welding failure sensitivity, heating wire disconnection fault, and heating wire aging fault. The first two have balanced fault data, while the latter three have imbalanced fault data. During the testing, the faulty sensor was placed in the sensor compartment and tested in a stable state. The operating conditions were normal atmospheric pressure, temperature was 18-25 ℃, and

moisture concentration was 40% -50%. The hydrogen concentration in the mixture was evenly distributed, and then entered the sensor array through a channel and was loaded onto the sensor array. The working and heating voltages of the sensing array were provided by a bidirectional adjustable power supply. The detection signals from the sensor array were collected in the data collection unit, which transmitted them to the computer system through the 232 bus. This application ran on Windows 11 and had a 3.2GHz CPU and 16GB of storage space. Figure 13 shows the performance comparison curves of different models in aging failure experiments.



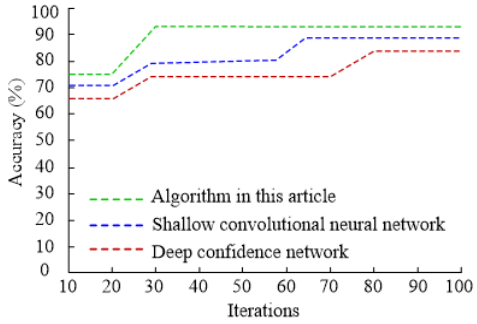
a) The recognition accuracy of three algorithms in aging failure.



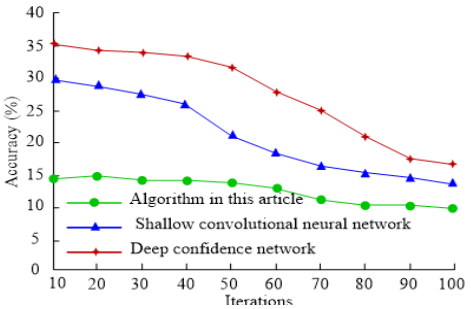
b) Loss function values of three algorithms on aging failure.

Figure 13. Loss value and accuracy curve of aging failure fault data diagnosed by different algorithms.

From Figure 13-a), the diagnostic accuracy curves obtained by the three algorithms through aging failure fault data testing showed consistency, all gradually increasing and tending to converge. As the iteration progressed, the algorithm proposed in the study achieved the greatest recognition accuracy of 96.8% when the number of iterations was 80. The shallow CNN achieved a second highest recognition accuracy of 89.1% when the number of iterations was 86. The deep confidence network achieved a minimum recognition accuracy of 80% when the number of iterations was 90. By comparison, the algorithm proposed in the study achieved the fastest convergence speed and the highest accuracy, which were improved by 7.7% and 16.8% respectively, compared to the latter. From Figure 13-b), the change trend of the loss function value curve of the proposed algorithm and the depth confidence network was similar, both were decline curves. However, the decline speed of the proposed algorithm was faster. When the number of iterations was 40, the corresponding curve finally got a stable loss function convergence value, which was 0.48%. The slope of the deep confidence network was smaller and the decline speed was slower. When the iteration times was 70, the curve began to converge, and the final value of the loss function was 30%. The difference between the two loss function values was an order of magnitude. The loss function curve of shallow CNN showed a small horizontal fluctuation first, and then a decreasing trend. This made the curve fail to obtain the convergence value of the loss function, and its minimum value was 3%. By comparison, the algorithm proposed by the research reduced it by 2.52%. To sum up, the data showed that the raised algorithm has improved the convergence speed and stability of the network because it integrated the generation of confrontation network and the improvement of loss function. Figure 14 shows the performance curves of different models in sensitivity failure experiments.



a) Identify accuracy of three algorithms in sensitivity failure.

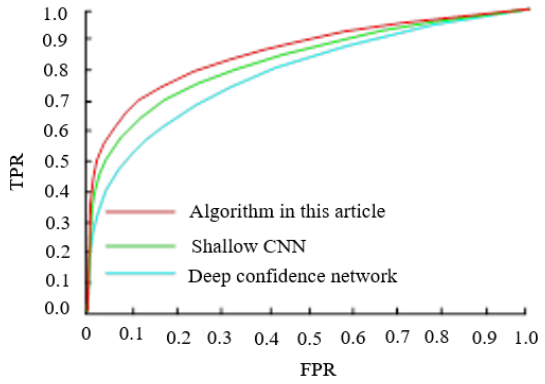


b) The loss function values of three algorithms in sensitivity failure.

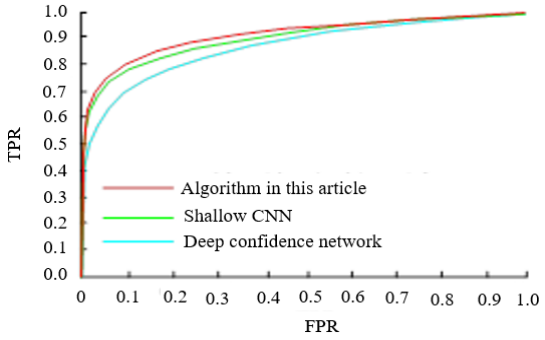
Figure 14. Loss value and accuracy curve of failure data with different algorithm diagnostic sensitivity.

From Figure 14-a), the proposed algorithm, shallow CNN, and deep confidence network began to converge at iterations of 30, 63, and 80, respectively. The convergence values of the recognition accuracy obtained by the three algorithms were 92.6%, 89.1%, and 82.5%, respectively. From this, the algorithm proposed in the study converged 33 and 50 iterations earlier than the latter two, and the recognition accuracy was improved by 3.5% and 10.1%, respectively. So this algorithm had a faster convergence speed and recognition effect. From Figure 14-b), both the shallow CNN and the deep confidence network showed a trend of slowly decreasing first and then rapidly decreasing. Among them, the value of the loss function of the shallow CNN decreased from 35%, and the minimum value of 16.5% was obtained when the number of

iterations was 100. That of the deep confidence network decreased from 30%, and the minimum value of 14% was obtained at the end of the iteration. The corresponding curve of the proposed algorithm fluctuated slowly around 15% at first. When the number of iterations was 60 and 80, the curve started to decline and converge, respectively, and finally the minimum value of 10% of the loss function was obtained. Figure 15 shows the Receiver Operating Characteristic (ROC) curves for aging failure and sensitivity failure diagnosis.



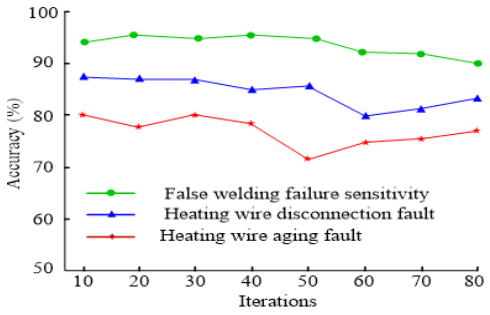
a) Aging failure.



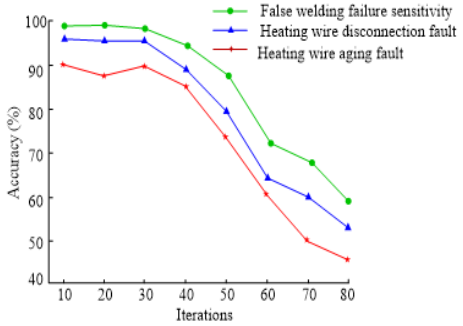
b) Sensitivity failure.

Figure 15. ROC curve for diagnosis of aging failure and sensitivity failure.

From Figure 15, in the diagnosis of aging failure and sensitivity failure, the Receiver Operating Characteristic Curve (ROC) composed of true and false positive rate obtained by the proposed algorithm performed better. Based on the y=x line, when diagnosing aging failure, the ROC curve of the proposed algorithm was farthest from the line. For example, when the false positive rate was 0.1, the true positive rates obtained by the three algorithms were 0.67, 0.61, and 0.51, respectively. Therefore, the proposed algorithm had a higher true positive rate. When the diagnostic sensitivity failed, the ROC curve obtained by the proposed algorithm was still further away from the y=x line. From this, the Area Under the Curve (AUC) value of the proposed model was larger, while the AUC of the other two models was smaller. Due to the fact that the ROC curve reflected the subject's discrimination ability, this indicated that the proposed algorithm had better recognition performance. Figure 16 shows the accuracy curves of diagnosing imbalanced fault data using different models.



a) Model in this artical.



b) Shallow convolutional network.

Figure 16. Accuracy curves for diagnosing imbalanced fault data using different models.

From Figure 16, the diagnostic accuracy curves of the model for three types of imbalanced fault data showed a small decrease during the iteration. Among them, the accuracy curve of false welding failure sensitive FD was generally stable during the convergence, while the heating wire disconnection fault and heating wire aging fault had a certain degree of fluctuation in the middle and later stages. However, the accuracy of all three was ultimately higher than 75%. The diagnostic accuracy of shallow CNNs for three types of imbalanced data showed a sharp decline trend, with the final accuracy being below 70%. By comparison, CNNs optimized using GAN performed better in small sample FD.

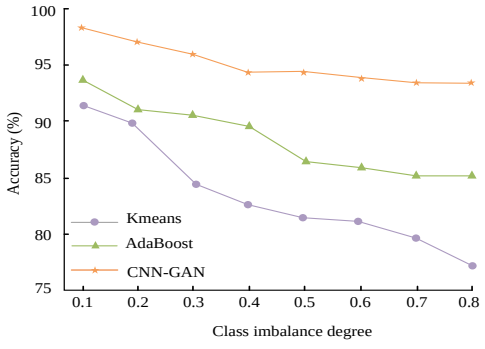


Figure 17. Model accuracy curves under different types of imbalances.

From Figure 17, as the class imbalance increased, the accuracy of the K-means, AdaBoost, and CNN-GAN models all showed a decreasing trend. Due to the fact that the K-means algorithm is only suitable for balanced data classification tasks, the accuracy of this model was relatively low. The other two models applicable to imbalanced data both achieved high convergence accuracy of 85.3% and 93.8%, respectively. The

algorithm proposed in the study had higher accuracy and better performance.

Table 2. Algorithm spatiotemporal complexity.

Model	Average run time(ms)	CPU usage rate (%)
Smote	562	5.6
AdaBoost	597	5.9
CNN-GAN	524	5.2

From Table 2, the spatiotemporal complexity of the three models was at a reasonable level. The average running time range of the algorithm was within 600ms, and the CPU utilization rate was less than 6%.

5. Conclusions

Hydrogen is a flammable and explosive gas. To avoid dangerous accidents, the use of HSs to monitor the concentration of hydrogen in real-time has become the key to the safe use of hydrogen. HSs are mostly made of semiconductor materials, which are greatly affected by environmental factors such as temperature. To accurately detect the operational status of sensors and timely diagnose their fault types, this study proposed a FD model based on deep convolutional GANs. The results showed that in the binary classification task, the proposed model had the smallest corresponding loss value, with a size of 2.0%, which was 0.8% and 1.2% less than the convergence loss values of the Densenet and ResNet models. In multi-classification tasks, the model still achieved a minimum loss value of 1.9%, which was a decrease of 0.6% and 2.0% compared to the latter two models. When diagnosing fault data with balanced distribution, the proposed model achieved the fastest convergence speed and highest diagnostic accuracy in aging failure diagnosis, with a convergence iteration number of 70 and a size of 96.8%, which was 7.7% and 16.8% higher than shallow CNNs and deep confidence networks, respectively. When diagnosing fault data with imbalanced distribution, the diagnostic accuracy of the models proposed in the study was higher than 75%, while the comparison model was lower than 70%. The above results indicated that the proposed model accelerated the convergence speed of accuracy and loss rate, overcame the need for a large amount of labeled training samples, and improved HSFD accuracy. With the increasing variety and volume of fault data, processing these data requires a lot of time. Subsequent research will consider GPU hardware factors in optimization to improve computing speed. In addition, the current research has not been able to extend the proposed method to other types of sensor faults, and it is expected to conduct in-depth research in the future. The limitations and challenges of this study lie in the fact that although using GAN to design a model solves the problem of data imbalance, the training process of the model is often difficult to control, which can hinder its application. Secondly, the model designed in this study has poor processing ability for high-dimensional and large-scale data. As a result, when the scale of the data to be processed is large, the training and prediction speed of the model is significantly slowed down. The future research directions of this study are as follows: The first direction is to attempt to further enhance the computational performance of the algorithm by improving its structure; the second direction is to test the application effect of the designed algorithm in fault detection of other gas sensors, and expand the applicability of the algorithm.

Competing Interests

The authors report there are no competing interests to declare.

References

[1] Abdullah Q., Ahmed A., Ali A., Yam F., Hassan Z., and Bououdina M., "Novel SnO_2-Coated β-Ga_2O_3 Nanostructures for Room Temperature Hydrogen Gas Sensor," *International Journal of Hydrogen Energy*, vol. 46, no. 9, pp. 7000-7010, 2021. https://doi.org/10.1016/j.ijhydene.2020.11.109

[2] Alajanbi M., Malerba D., and Liu H., "Distributed Reduced Convolution Neural Networks" *Mesopotamian Journal of Big Data*, vol. 2021, no. 2021, pp. 25-28, 2021. https://doi.org/10.58496/MJBD/2021/005

[3] Antar A., Ahmed M., and Ahad M., "Challenges in Sensor-Based Human Activity Recognition and a Comparative Analysis of Benchmark Datasets: A Review," *in Proceedings of the Joint 8th International Conference on Informatics, Electronics and Vision and the 3rd International Conference on Imaging, Vision and Pattern Recognition*, Spokane, pp. 134-139, 2019. https://doi.org/10.1109/ICIEV.2019.8858508

[4] Calado A., Soares F., and Matos D., "A Review on Commercially Available Anthropomorphic Myoelectric Prosthetic Hands, Pattern-Recognition-Based Microcontrollers and sEMG Sensors Used for Prosthetic Control," *in Proceedings of the IEEE International Conference on Autonomous Robot Systems and Competitions*, Porto, pp. 1-6, 2019. https://doi.org/10.1109/ICARSC.2019.873362

[5] Chen Z., Cen J., and Xiong J., "Rolling Bearing FD Using Time-Frequency Analysis and Deep Transfer Convolutional Neural Network," *IEEE Access*, vol. 8, pp. 150248-150261, 2020. https://doi.org/10.1109/ACCESS.2020.3016888

[6] Choudhuri S., Adeniye S., and Sen A., "Distribution Alignment Using Complement Entropy Objective and Adaptive Consensus-Based Label Refinement for Partial Domain Adaptation," *Artificial Intelligence and*

Applications, vol. 1, no. 1, pp. 43-51, 2023. https://doi.org/10.47852/bonviewAIA2202524

[7] Daihong J., Sai Z., Lei D., and Yueming D., "Multi-Scale Generative Adversarial Network for Image Super-Resolution," *Soft Computing*, vol. 26, no. 8, pp. 3631-3641, 2022. https://doi.org/10.1007/s00500-022-06822-5

[8] Darvishi H., Ciuonzo D., Eide E., and Rossi P., "Sensor-Fault Detection, Isolation and Accommodation for Digital Twins via Modular Data-Driven Architecture," *IEEE Sensors Journal*, vol. 21, no. 4, pp. 4827-4838, 2020. https://doi.org/10.1109/JSEN.2020.3029459

[9] Dasebenezer G. and Joselin B., "TSO Clustered Protocol to Extend Lifetime of IoT Based Mobile Wireless Sensor Networks," *The International Arab Journal of Information Technology*, vol. 20, no. 4, pp. 559-566, 2023. https://doi.org/10.34028/iajit/20/4/1

[10] Doostmohammadian M. and Meskin N., "Sensor Fault Detection and Isolation via Networked Estimation: Full-Rank Dynamical Systems," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 2, pp. 987-996, 2021. https://doi.org/10.1109/TCNS.2020.3029165

[11] Duan L., Guo L., Gao H., Wu X., and Dong X., "Deep Focus Parallel Convolutional Neural Network for Imbalanced Classification of Machinery Fault Diagnostics," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 11, pp. 8680-8689, 2020. https://doi.org/10.1109/TIM.2020.2998233

[12] Fang B., Jiang M., Shen J., and Stenger B., "Deep Generative Inpainting with Comparative Sample Augmentation," *Journal of Computational and Cognitive Engineering*, vol. 1, no. 4, pp. 174-180, 2022. https://doi.org/10.47852/bonviewJCCE2202319

[13] Gao S., Pei Z., Zhang Y., and Li T., "Bearing FD Based on Adaptive Convolutional Neural Network with Nesterov Momentum," *IEEE Sensors Journal*, vol. 21, no. 7, pp. 9268-9276, 2021. https://doi.org/10.1109/JSEN.2021.3050461

[14] Goldthau A. and Tagliapietra S., "Energy Crisis: Five Questions that must be Answered in 2023," *Nature*, vol. 612, no. 7941, pp. 627-630, 2022. https://doi.org/10.1038/d41586-022-04467-w

[15] Guo M., Yang N., and Chen W., "Deep-Learning-Based Fault Classification Using Hilbert-Huang Transform and Convolutional Neural Network in Power Distribution Systems," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 6905-6913, 2019. https://doi.org/10.1109/JSEN.2019.2913006

[16] Guo Y., Mustafaoglu Z., and Koundal D., "Spam Detection Using Bidirectional Transformers and Machine Learning Classifier Algorithms," *Journal of Computational and Cognitive Engineering*, vol. 2, no. 1, pp. 5-9, 2022. https://doi.org/10.47852/bonviewJCCE2202192

[17] Hasan M., Abir F., Siam M., and Shin J., "Gait Recognition with Wearable Sensors Using Modified Residual Block-Based Lightweight CNN," *IEEE Access*, vol. 10, pp. 42577-42588, 2022. https://doi.org/10.1109/ACCESS.2022.3168019

[18] Jeong K., Choi S., and Choi H., "Sensor Fault Detection and Isolation Using a Support Vector Machine for Vehicle Suspension Systems," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 4, pp. 3852-3863, 2020. https://doi.org/10.1109/TVT.2020.2977353

[19] Jiang Q., Chang F., and Sheng B., "Bearing Fault Classification Based on Convolutional Neural Network in Noise Environment," *IEEE Access*, vol. 7, pp. 69795-69807, 2019. https://doi.org/10.1109/ACCESS.2019.2919126

[20] Kar M., Neog D., and Nath M., "Retinal Vessel Segmentation Using Multi-Scale Residual Convolutional Neural Network (MSR-Net) Combined with Generative Adversarial Networks," *Circuits, Systems, and Signal Processing*, vol. 42, no. 2, pp. 1206-1235, 2023. https://doi.org/10.1007/s00034-022-02190-5

[21] Kench S. and Cooper S., "Generating Three-Dimensional Structures from a Two-Dimensional Slice with Generative Adversarial Network-Based Dimensionality Expansion," *Nature Machine Intelligence*, vol. 3, no. 4, pp. 299-305, 2021. https://doi.org/10.1038/s42256-021-00322-1

[22] Kim M., Yun J., and Park P., "An Explainable Convolutional Neural Network for FD in Linear Motion Guide," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4036-4045, 2021. https://doi.org/10.1109/TII.2020.3012989

[23] Kiranyaz S., Ince T., Abdeljaber O., Avci O., and Gabbouj M., "1-D Convolutional Neural Networks for Signal Processing Applications," *in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, pp. 8360-8364, 2019. https://doi.org/10.1109/ICASSP.2019.8682194

[24] Lei X., Sun L., and Xia Y., "Lost Data Reconstruction for Structural Health Monitoring Using Deep Convolutional Generative Adversarial Networks," *Structural Health Monitoring*, vol. 20, no. 4, pp. 2069-2087, 2021. https://doi.org/10.1177/1475921720959226

[25] Li J., Wang Y., Wang P., Bai Q., Gao Y., Zhang H., and Jin B., "Pattern Recognition for Distributed Optical Fiber Vibration Sensing: A Review," *IEEE Sensors Journal*, vol. 21, no. 10, pp. 11983-11998, 2021. https://doi.org/10.1109/JSEN.2021.3066037

[26] Li Y., "Research and Application of Deep Learning in Image Recognition," *IEEE 2nd*

International Conference on Power, Electronics and Computer Applications, Shenyang, pp. 994-999, 2022. https://doi.org/10.1109/ICPECA53709.2022.9718847

[27] Lindsay G., "Convolutional Neural Networks as a Model of the Visual System: Past, Present, and Future," *Journal of Cognitive Neuroscience*, vol. 33, no. 10, pp. 2017-2031, 2021. https://doi.org/10.1162/jocn_a_01544

[28] Liu R., Wang F., Yang B., and Qin S., "Multiscale Kernel Based Residual Convolutional Neural Network for Motor FD under Nonstationary Conditions," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 6, pp. 3797-3806, 2020. https://doi.org/10.1109/TII.2019.2941868

[29] Lyu C., Huo Z., Cheng X., Jiang J., Alimasi A., and Liu H., "Distributed Optical Fiber Sensing Intrusion Pattern Recognition Based on GAF and CNN," *Journal of Lightwave Technology*, vol. 38, no. 15, pp. 4174-4182, 2020. https://doi.org/10.1109/JLT.2020.2985746

[30] Mishra G., Kumar D., Chaudhary V., and Kumar S., "Design and Sensitivity Improvement of Microstructured-Core Photonic Crystal Fiber Based Sensor for Methane and Hydrogen Fluoride Detection," *IEEE Sensors Journal*, vol. 22, no. 2, pp. 1265-1272, 2021. https://doi.org/10.1109/JSEN.2021.3131694

[31] Sharma S., Chhetry A., Zhang S., Yoon H., Park C., and Kim H., "Hydrogen-Bond-Triggered Hybrid Nanofibrous Membrane-Based Wearable Pressure Sensor with Ultrahigh Sensitivity over a Broad Pressure Range," *ACS Nano*, vol. 15, no. 3, pp. 4380-4393, 2021. https://doi.org/10.1021/acsnano.0c07847

[32] Singla M., Sharma D., and Jaggi N., "Effect of Transition Metal (Cu and Pt) Doping/Co-Doping on Hydrogen Gas Sensing Capability of Graphene: A DFT Study," *International Journal of Hydrogen Energy*, vol. 46, no. 29, pp. 16188-16201, 2021. https://doi.org/10.1016/j.ijhydene.2021.02.004

[33] Tan Y., Wang K., Su X., Xue F., and Shi P., "Event-Triggered Fuzzy Filtering for Networked Systems with Application to Sensor Fault Detection," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 6, pp. 1409-1422, 2020. https://doi.org/10.1109/TFUZZ.2020.2977252

[34] Tigrini A., Pettinari L., Verdini F., Fioretti S., and Mengarelli A., "Shoulder Motion Intention Detection through Myoelectric Pattern Recognition," *IEEE Sensors Letters*, vol. 5, no. 8, pp. 1-4, 2021. https://doi.org/10.1109/LSENS.2021.3100607

[35] Wang J., Wang D., Wang S., Li W., and Song K., "Fault Diagnosis of Bearings Based on Multi-Sensor Information Fusion and 2D Convolutional Neural Network," *IEEE Access*, vol. 9, pp. 23717-23725, 2021. https://ieeexplore.ieee.org/abstract/document/9345801

[36] Wen L., Li X., and Gao L., "A New Reinforcement Learning Based Learning Rate Scheduler for Convolutional Neural Network in Fault Classification," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 12, pp. 12890-12900, 2021. https://doi.org/10.1109/TIE.2020.3044808

[37] Yang X., Yan J., Chen Z., Ding H., and Liu H., "A Proportional Pattern Recognition Control Scheme for Wearable A-Mode Ultrasound Sensing," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 1, pp. 800-808, 2020. https://doi.org/10.1109/TIE.2019.2898614

[38] Zhang J., Christofides P., He X., Albalawi F., Zhao Y., and Zhou D., "Intermittent Sensor Fault Detection for Stochastic LTV Systems with Parameter Uncertainty and Limited Resolution," *International Journal of Control*, vol. 93, no. 4, pp. 788-796, 2020. https://doi.org/10.1080/00207179.2018.1490819

[39] Zhu S., Guendel R., Yarovoy A., and Fioranelli F., "Continuous Human Activity Recognition with Distributed Radar Sensor Networks and CNN-RNN Architectures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-15, 2022. https://doi.org/10.1109/TGRS.2022.3189746

Jun Wang was born in Jiu Jiang City, JiangXi Province, China in 1982. He received his B.Sc. degree in 2005 and M.Sc. degree in 2009 from South-central University for Nationality, studying for a PhD in Tongji University in 2020. His main research interests include Advanced Sensors, Artificial Intelligence Algorithms, Detection Technology and Automation Devices. Jun Wang, a member of the Natural Computing and Digital Smart City Professional Committee of the Chinese Society for Artificial Intelligence, has invented a highly sensitive multi-mode hydrogen sensor and put it into market use.

Yuanxi Wang was born in Shanghai, China in 1993. He received his B.Sc. degree in 2016 from Shanghai University of Electric Power, received his M.Sc. degree in 2018 from University of Hull in the United Kingdom. Studying at Tongji University in 2020. His main research direction is the preparation of solid Electrolytes and New Electrochemical Sensors.