APPLIED RESEARCH

A New Method for Diagnosing Motor Bearing Faults Based on Gramian Angular Field Image Coding and Improved CNN-ELM

YUHAN ZHANG [ID]1, LEI SHANG1, HAIBO GAO [ID]1, YELAN HE2, XIAOBIN XU3, AND YAJIE CHEN4

1Ocean and Energy Power Engineering Department, Wuhan University of Technology, Wuhan 430063, China
2School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan 430063, China
3School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China
4Shanghai Marine Diesel Engine Research Institute, Shanghai 201108, China

Corresponding authors: Yuhan Zhang (yh_zhang@whut.edu.cn) and Haibo Gao (hbgao_whut@126.com)

ABSTRACT The most common cause of electric motor failure is the bearings, and so methods for fast and accurate diagnosis of motor bearing failure are urgently needed. Traditional fault diagnosis methods have high uncertainty and complexity since they require manual extraction of features. Deep learning has shown good performance in electrical equipment fault detection, and it can directly complete end-to-end diagnosis of motor faults, avoiding human involvement. Here, a new fault diagnosis method is presented which combines Gramian angular field (GAF) image coding, extreme learning machine (ELM) and convolutional neural network (CNN). The method has three main stages: First of all, GAF is utilized to convert the acquired vibration break signals into 2-D pictures. Next, the enhanced CNN model is taken to identify the elements of the converted image quickly and accurately. Finally, the ELM is used as the final classifier to gain further accuracy and diagnostic speed of fault classification. Experiments were designed to validate the proposed method using two different motor bearing fault datasets at Case Western Reserve University and autonomous experiment and performance is compared with several commonly used intelligent diagnosis algorithms. The proposed method's accuracy in the experiment designed in this paper can reach 99.2% at most, and it only takes 0.835s to complete the diagnosis, which outperforms traditional diagnostic methods on both datasets and improving the maximum diagnostic accuracy by 33.6%. The findings indicate that this method can classify various fault types efficaciously, and has the benefits of quick diagnosis, high accuracy, and good generalization ability.

INDEX TERMS Gramian angular field image coding, convolutional neural network, extreme learning machine, fault diagnosis.

I. INTRODUCTION

Since Faraday's first demonstration of electrical energy conversion into mechanical energy in 1821 the electric motor has evolved into a widely used form of power in daily use in both domestic and industrial situations. As the time of use

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

of an electric motor increases failures can occur which can damage the equipment itself, bring about economic losses, or contribute to major accidents and casualties. The majority (45%-55%) of motor failures are related to their bearings [1], which have as a result been the focus of much research into reliable fault detection and diagnosis [2].

The development of increasingly sophisticated fault diagnosis technology has required a multidisciplinary approach.

Data-driven methods [3] and analytical modelling methods [4] are two commonly used fault diagnosis methods. The analytical model method requires the establishment of a mathematical model to depict the mechanism of the fault. This method is more difficult to establish as the complexity of the system increases and the accuracy of the model cannot be guaranteed. In addition, established models are less generally applicable and cannot solve the same fault problems for different motors, so analytical model methods are not often used in practice.

Due to the rapid development of the IoT (Internet of Things), new technologies, sophisticated sensors, and data collection systems are now applied to motor bearings [5]. Monitoring vibration data, current and other signals at different locations and times is becoming easier and more cost efficient, so technicians can carry out fault diagnosis of motor bearings by mining and analyzing huge amounts of historical data available [6]. The manual extraction of features from such large data sets in challenging and may cause the loss of original data. Besides, the inability to correctly select the extracted features will also affect the final classification results, therefore, against the background of intelligent industrial manufacturing in the ''big data'' era, it is especially critical to establish a data-driven intelligent fault diagnosis method.

During the past few years, artificial intelligence has developed rapidly, various industries are competing to combine artificial intelligence and fault diagnosis methods. Unlike classical fault diagnosis methods that rely on manual signal analysis, intelligent diagnosis algorithms can directly extract beneficial properties from massive amounts of data [7]. The four steps below make up a typical intelligent diagnosis method: gathering data, extracting features, choosing features, and identifying faults [8]. Feature extraction is the signal processing of the raw data collected by the sensor to extract a useful set of features for fault type identification. Wavelet transform (WT), Fourier transform (FFT), and empirical modal decomposition (EMD) are the most popular feature extraction techniques [9], [11]. The extracted features are usually high dimensional, and there may also be invalid information affecting the final diagnosis results. The extracted features need to be further filtered by feature selection, which can enhance the precision and computational effectiveness of the diagnosis results. Common methods used for this step are principal component analysis (PCA) [12] and independent component analysis (ICA) [13]. The further filtered features are subsequently placed into a shallow machine learning model for training. Typical shallow models include artificial neural networks (ANN), support vector machines (SVM) and k-nearest neighbors (KNN) [14] and the training determines the model's parameters to complete fault classification. Despite the positive outcomes of the aforementioned conventional intelligent diagnosis methods, there remain some problems as follows: (1) the different real world industrial environments may require different feature selection methods, and appropriate feature selection methods are highly dependent on situation specific expertise; (2) most of the extracted features are superficial features, which are not generally applicable in the face of complex classification problems; and (3) limited by mechanical systems' physical properties, changes in fault conditions can significantly affect the evaluation criteria of feature extraction and also affect the mining of new features.

The theory of deep learning was put forward by Hinton et al. in 2006 [15], which is essentially a neural network with numerous hidden layers, extracting deeper features of data through deep network structures. With the popularity of GPUs, the computing power of computers has greatly increased, and several fields have achieved success with deep learning, including computer vision [16], driverless cars [17] and natural language processing [18]. Its powerful data processing capability can extract deeper special characteristics from data, which can effectively compensate for the weakness of traditional intelligent diagnosis. Since 2013 when Tamilselvan and Wang have used the deep belief network (DBN) to diagnose problems with aeroplane engines and power transformers, research on deep learning models for fault diagnosis has continued to intensify [19]. Because of their powerful feature extraction capability, convolutional neural networks (CNNs) have been extensively utilized in fault diagnosis over the past few years. David et al. converted bearing vibration signals into time-frequency pictures integrated with CNN based fault diagnosis [20]. Wen et al. transformed bearing vibration signals into grayscale maps combined with CNN to complete the diagnosis. Gong et al. transmogrified the signal into a grayscale map and used global average pooling in place of the conventional CNN approach of using a fully connected layer to decrease the number of parameters, combined with an SVM classifier to complete the diagnosis [21]. Han et al. proposed a method to solve the problem of out-of-distribution (OOD) of fault samples using deep neural network integration [22]. To complete fault classification, Chen et al. transformed the signal data into a time-frequency map and combined CNN with a square pooling structure and an extreme learning machine (ELM) [23]. Han et al. presented a method for converting the signal for Gramian angular field (GAF) image coding into an image combined with a capsule network for diagnosing bearing faults [24]. Inspired by the previous work and combined with the requirement of fast and accurate fault diagnosis with minimal human influence, this study presents a novel method which is the first to combine the GAF-CNN and ELM aspects. The main insights and contributions of this article can be summarized as follows:

(1) This paper establishes GAF as the optimal image coding method through comparative experiments, effectively establishing a bridge between computer vision and fault diagnosis.

(2) To address the problem that the parameters of the classical network model are heavily concentrated in the fully

connected layer, the use of a global mean pooling layer is proposed instead of a fully connected layer, and a batch regularization layer is added after each convolution operation. The final model not only has greatly reduced parameters, but also has enhanced generalization ability, and the superiority of the improved model is demonstrated through validation on different experimental data sets.

(3) The traditional Softmax classifier is replaced by ELM, which further improves the diagnosis speed and accuracy, and is then combined with GAF and improved CNN model shape to form a unique GAF-CNN-ELM (GCE) fault diagnosis method.

(4) Separate method validation on the Case Western Reserve University [25]dataset and the autonomous experimental dataset was conducted, and the reliability of the results was further enhanced by combining the publicly available dataset and the autonomous experimental dataset.

In the remainder of the paper, the following structure is followed. In Section II, the principles of GAF image coding, CNNs and extreme learning machines are described. The improvements made by the CNN model and the overall structure of the proposed method which are displayed in Section III. In Section IV, experiments are presented which validate the proposed fault diagnosis method using different datasets. Lastly, Section V contains the conclusions.

II. PRINCIPLE OF RELATED ALGORITHMS

A. THE PRINCIPLE OF GRAMIAN ANGULAR FIELD (GAF)

The rapid development of artificial intelligence within the field of computer vision technology, suggested the possibility of combining motor fault diagnosis with computer vision technologies. To connect the two fields, the key is to seek to convert motor fault signals into images recognizable using computer vision. Wang et al. initially proposed a GAF algorithm, which encodes 1-D time series data into unique 2-D pictures [26]. The specific implementation steps are outlined below:

The time series $X = \{x_1, x_2, \ldots, x_n\}$ containing measured values are normalized to values between 0 and 1:

$$\tilde{x} = \frac{x_i - min(X)}{max(X) - min(X)}, (i = 1, 2, \ldots n) \quad (1)$$

Then, the one-dimensional time series are transformed into the polar coordinate system using the arccos() function and noting the time stamp as the radius, as demonstrated by Equation (2)

$$\begin{cases} \phi = \arccos(\tilde{x}_i), 0 \leq \tilde{x}_i \leq 1, & \tilde{x}_i \in \tilde{X} \\ r = \frac{t_i}{N}, & t_i \in N \end{cases} \quad (2)$$

where t_i denotes the time stamp, N is a constant factor to adjust the range of polar coordinates, and ϕ is the angular cosine of each value in the time series.

From equation (2), we can see that the range of values of ϕ is $\left[0, \frac{\pi}{2}\right]$, and the cosine of the angle in this range decreases monotonically as $\tilde{x}_i$ increases, and the different ϕ
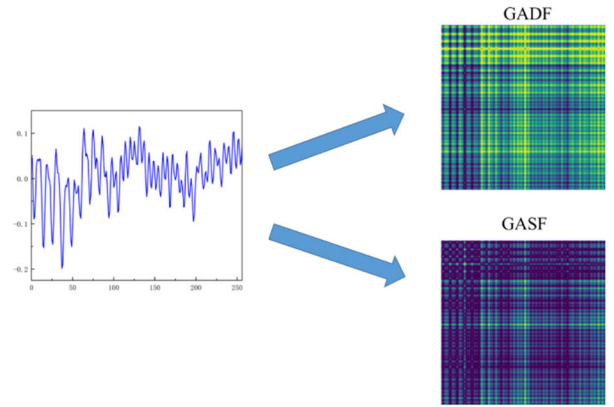

FIGURE 1. GADF/GASF mapping diagram.

values produce corresponding distortions between different angular points on the polar circle.

Finally, after transforming the scaled time series to values in the polar coordinate system, the angular perspective view is used by taking into account the triangular sum or difference between each point to determine time-series correlations at different time intervals. GAF has two different encoding methods, Gram's Angular Sum Field (GASF) and Gram's Angular Difference Field (GADF), which are as shown in Equation (3) and Equation (4):

$$GASF = \begin{bmatrix} \cos(\phi_1 + \phi_1) & \ldots & \cos(\phi_1 + \phi_n) \\ \vdots & \ddots & \vdots \\ \cos(\phi_n + \phi_1) & \ldots & \cos(\phi_n + \phi_n) \end{bmatrix} \quad (3)$$

$$GADF = \begin{bmatrix} \sin(\phi_1 - \phi_1) & \ldots & \sin(\phi_1 - \phi_n) \\ \vdots & \ddots & \vdots \\ \sin(\phi_n - \phi_1) & \ldots & \sin(\phi_n - \phi_n) \end{bmatrix} \quad (4)$$

The use of Gramian angular field method can preserve the time-series characteristics of the signal, because the time series transformed by polar coordinates are input from the upper left corner to the lower right corner accompanied by an increase in time, so that the converted image can maximize the preservation of the features possessed by the original signal, which can effectively help us to identify the fault types using convolutional neural networks. Fig. 1 depicts the outcome of using two different GAF coding techniques to transform a time-series signal into a picture.

B. THE PRINCIPLE OF CONVOLUTIONAL NEURAL NETWORK (CNN)

CNN have feed-forward structure, which can automatically extract local features without human intervention when classifying images. As an efficient and classical deep learning algorithm, it has been broadly utilized in many different fields requiring pattern classification. The convolutional, pooling and fully connected layers make up the traditional CNN architecture, and this paper introduces a batch normalization layer and substitutes the fully connected layer with a global

average pooling layer [27], [28]. The mathematical model for each layer of the CNN architecture is explained in this section.

1) CONVOLUTIONAL LAYER

The convolutional layer is the core of all types of CNN models, which is used to slide the convolutional kernels over the input feature map to extract the structural features hidden inside the data by convolutional operations on the data within the local perceptual field. Equal steps are taken to traverse the input feature map by the same convolution kernel. Weight sharing is what is used in this process, which significantly lowers the amount of network parameters in the convolution layer and reduces the memory required for the computation process. The result of the convolution operation is obtained as the output after the activation function, and Equation (5) provides the mathematical expression for the convolution layer:

$$x_j^l = \mathbf{f}(\sum_{i \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l) \tag{5}$$

where $(*)$ is the convolution operator, M_j is the set of input feature vectors, x_j^l denotes the jth feature graph of the lth layer, k_{ij}^l is the weight matrix of the convolution kernel, b_j^l is the additive deviation of the current feature map, and $\mathbf{f}(\cdot)$ is the nonlinear activation function.

2) POOLING LAYER

The pooling layer then follows the convolutional layer. Its primary function is to further decrease the amount of network parameters by reducing the dimensionality of the feature map following the convolutional operation. The pooling layer operates similarly to the convolutional layer by traversing the feature map with a sliding window and using the statistics of the sliding region as the sampling value, whose mathematical expression is shown in Equation (6)

$$x_j^l = \mathbf{f}\left(\beta_j^l \mathbf{down}\left(x_j^{l-1} + b_j^l\right)\right) \tag{6}$$

where β_j^l is the multiplicative deviation, b_j^l is the additive deviation, and $\mathbf{down}(\cdot)$ is the pooling function.

The most prevalent pooling functions are maximum pooling and average pooling. The former computes the maximum value in the sliding region as the output, while the latter computes the average value in the sliding region as the output.

3) GLOBAL AVERAGE POOLING LAYER

After the convolution and pooling operations, rather than using the traditional fully connected layer, the global average pooling layer is used. This layer does not need to set parameters, which can decrease the amount of network parameters and prevent overfitting, and its operation is to average all channels of the final output feature map, and the
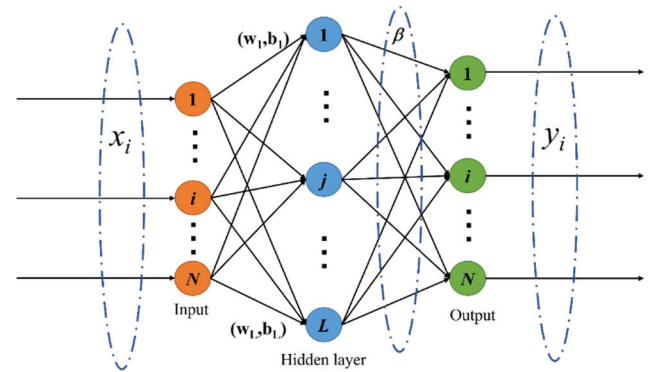

FIGURE 2. Structure of ELM.

mathematical formula is given in Equation (7):

$$S_{avg}^l = \frac{1}{n}\sum_{i=1}^{n} X_i^l \tag{7}$$

where S_{avg}^l denotes the result of global mean pooling of the lth layer feature map, X_i^l denotes the pixel points corresponding to the mean pooling kernel, and n denotes the total amount of all pixel points in the pooling kernel.

4) BATCH NORMALIZATION LAYER

Batch normalization (BN) is a method for optimizing networks proposed by Ioffe and Szegedy [29]. Its computation of mean and variance estimates on a small batch of training sets can reduce the internal covariate transfer, accelerate the convergence, and improve the network's generalization ability. The mathematical expression for the operation of the BN layer is shown in Equation (8):

$$\begin{cases} \mu_{BN} = \frac{1}{N}\sum_{i=1}^{N} a_i \\ \delta_{BN}^2 = \frac{1}{N}\sum_{i=1}^{N}(a_i - \mu_{BN})^2 \\ \widehat{a_i} = \frac{a_i - \mu_{BN}}{\sqrt{\delta_{BN}^2 + \xi}} \\ b_i = \gamma\widehat{a_i} + \beta \end{cases} \tag{8}$$

where N denotes the size of the training set for small batches, a_i is the input of the BN layer, b_i is the output of the BN layer, ξ is a constant, μ_{BN} is the mean of the input, δ_{BN}^2 is the variance of the input, $\widehat{a_i}$ is the normalized value of the input, and γ and β are two learnable parameters.

C. THE PRINCIPLE OF EXTREME LEARNING MACHINE (ELM)

Huang et al. proposed a single hidden layer feed-forward neural network called the extreme learning machine [30]. Unlike the traditional SLFN algorithm, the weights and biases of the ELM hidden layer are randomly generated and remain constant during training. After determining the amount of

nodes in the hidden layer, the connection matrix of the hidden layer and the output layer can be derived through the generalized inverse matrix.

Figure. 2 depicts the ELM structure, given a training set $\{x_i, y_i \mid x_i \in R^{D \times N}, y_i \in R^{1 \times N}, i = 1, 2, \ldots N\}$, x_i stands for the ith input, y_i is the label corresponding to the ith input, the amount of nodes in the hidden layer is L, and the output of the hidden layer is noted as $h(x)$, which is shown as follows:

$$h(x) = [h_1(x), \ldots, h_L(x)] \tag{9}$$

$h(x)$ denotes the non-linear mapping of ELM, and $h_j(x)$ denotes the output of the jth hidden layer node, which is determined as illustrated in Equation (10):

$$h_j(x) = g(w_j, b_j, x) = g(w_j x + b_j), w_j \in R^{N \times L}, b_j \in R^{1 \times L} \tag{10}$$

where $g(\cdot)$ is the activation function, The weight and deviation are denoted by w_j and b_j, which are produced at random by the jth hidden node.

The output of ELM is displayed below after the hidden layer.

$$f_L(x) = \sum_{j=1}^{L} \beta_j h_j(x) = h(x)\beta \tag{11}$$

where $\beta = [\beta_1, \ldots \beta_L]^T$ is the weight between the hidden layer and the output layer.

w_j and b_j are determined by randomization and the hidden layer output H is shown as follows:

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} h_1(x_1) & \ldots & h_L(x_1) \\ & \ldots & \\ h_1(x_N) & \ldots & h_L(x_N) \end{bmatrix} \tag{12}$$

Now only the output layer weight β is required to be solved, and the network output is $H\beta$, using $H\beta$ and the sample label $Y = [y_1, \ldots y_N]^T$ to determine the minimized squared difference as the evaluation training error (objective function), with the goal of identifying the best solution by having the smallest objective function, which is as follows:

$$\min \|H\beta - Y\|^2, \beta \in R^{L \times N} \tag{13}$$

The optimal solution β^* for the output weights is finally calculated as shown in the following equation:

$$\beta^* = H^+ Y \tag{14}$$

where H^+ is the Moore-Penrose generalized inverse matrix of the matrix H.

Compared with traditional networks, ELM reduces its training parameters by randomly generating hidden layer weights and biases, which increases learning speed and increases the generalization capacity of the network by minimizing the number of weights and training errors.
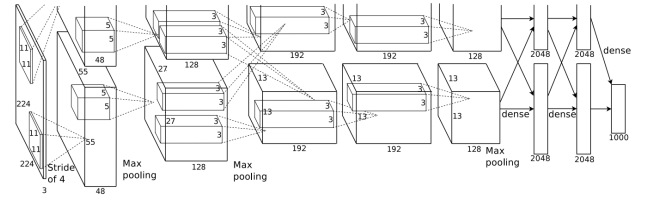

FIGURE 3. Structure of AlexNet.

TABLE 1. The main structural parameters of the alexnet model.

Layer Name	Kernel Num	Kernel Size	Output Shape	Parameters
Conv-1	96	11×11	55×55×96	34,944
Maxpool-1	96	3×3	27×27×96	0
Conv-2	256	5×5	27×27×256	614,656
Maxpool-2	256	3×3	13×13×256	0
Conv-3	384	3×3	13×13×384	885,120
Conv-4	384	3×3	13×13×384	1,327,488
Conv-5	256	3×3	13×13×256	884,992
Maxpool-2	256	3×3	6×6×256	0
FC-1			4096×1	37,752,832
FC-2			4096×1	16,781,312
FC-3			1000×1	4,097,000
Total				**62,378,344**

III. THE PROPOSED GAF-CNN-ELM METHOD

Throughout this section, a brand-new fault diagnosis technique based on the combined use of GAF, CNN and ELM, called GCE is proposed. The method is capable of extracting the feature information inherent in the fault signal, and also offers fast training speed and good generalization capability. GCE is performed as follows (1) 1-D signal is transformed into 2-D graph input (2) improved CNN is used for feature extraction (3) ELM is used as a classifier.

A. IMPROVED CNN STRUCTURE

Theoretically, the more complex the network structure is, the higher the classification accuracy, but the corresponding computation time will be longer. Considering that bearing fault diagnosis should be highly accurate as well as rapid so that faults can be dealt with promptly to avoid more serious losses. Application methods from the field of computer vision uses overly complex network models such as ResNet101 and GoogleNet but are not applicable because their network model is too complex, the diagnosis efficiency is not high, so this paper suggests an enhanced CNN network model based on the AlexNet network [31] with a relatively simple model but with performance improved over the earliest LeNet-5 network [32] for fault diagnosis.

The AlexNet network [31] was designed by Alex Krizhevshy et al. in 2012 in Geoff Hinton's lab at the University of Toronto, and Figure. 3 depicts its structure. The AlexNet network structure is divided into two levels, each containing 11 more layers, containing three fully connected layers, three maximum pooling layers, and five convolutional layers.

From Table 1, it can be seen that the AlexNet network has a substantial number of parameters, most of which are
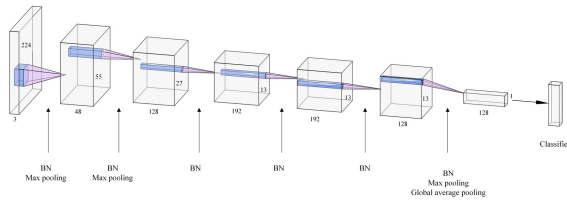
FIGURE 4. Architecture of improved CNN model.

TABLE 2. The main structural parameters of the improved CNN model.

Layer Name	Kernel Num	Kernel Size	Output Shape	Parameters
Conv-1	48	11×11	55×55×48	17,472
BN-1			55×55×48	96
Maxpool-1	96	3×3	27×27×48	0
Conv-2	128	5×5	27×27×128	153,728
BN-2			27×27×128	256
Maxpool-2	128	3×3	13×13×128	0
Conv-3	192	3×3	13×13×192	221,376
BN-3			13×13×192	384
Conv-4	192	3×3	13×13×192	331,968
BN-4			13×13×192	384
Conv-5	128	3×3	13×13×128	221,312
BN-5			13×13×128	256
Maxpool-2	128	3×3	6×6×128	0
GAP	128	6×6	1×1×128	
Total				947,232

located at the fully connected layer. Therefore, considering that there are few types of faults related to motor bearing fault diagnosis and it needs to complete diagnosis quickly, the improved CNN network model that is suggested in this paper retains the same structure of the basic feature extraction layer of the original model, and due to the symmetry of AlexNet model, only the upper half level of the original model is taken to decrease the number of training parameters, and the fully connected layer is replaced with a global average pooling layer. To boost the model's capacity for generalization and accelerate convergence during training, each convolutional layer in the improved network model is followed by a batch normalization layer, and the improved CNN structure is depicted in Figure. 4.

Consequently, the number of parameters in the proposed CNN network model is 98% less than that of the original AlexNet network model resulting in less required computing power and faster network training speeds. The specific performance will be further characterised in the following experiments.

B. GCE(GAF-CNN-ELM)

After determining the structure of the CNN, a new combined method of GAF-CNN-ELM is proposed for motor fault diagnosis. There are two different GAF encoding methods, GASF and GADF, and due to the necessity to reflect the effectiveness of the selected encoding method, the article also explores the encoding methods of the grayscale map and the time-frequency map to process the fault signal at
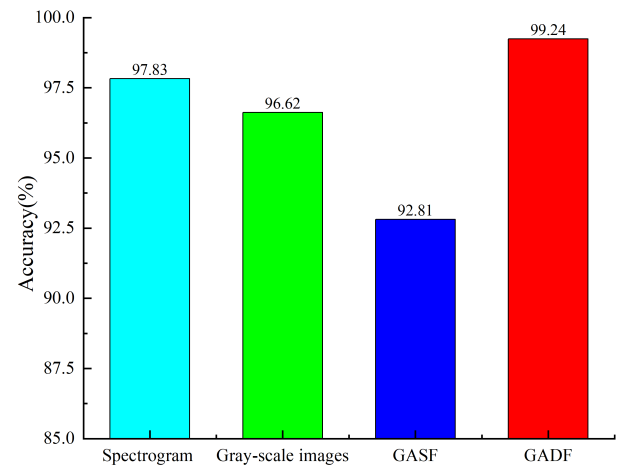

FIGURE 5. Accuracy of the different coding methods.
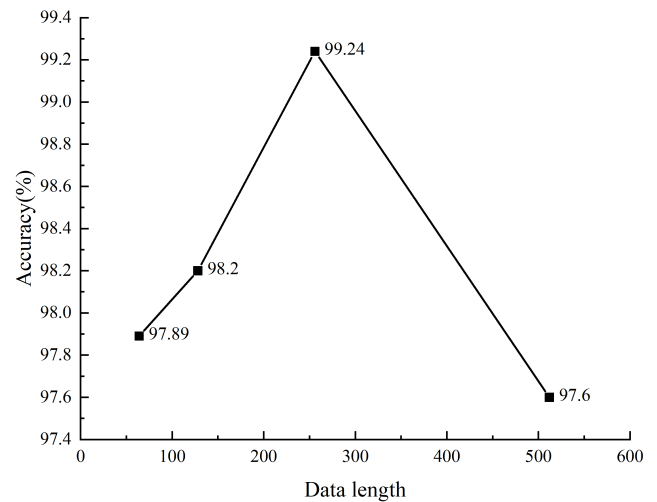

FIGURE 6. Accuracy of the different data length.

the same time. These encoding methods are all combined with the improved CNN model before the experimental stage, (Figure. 5). The accuracy of GADF is the highest when performing pattern recognition, therefore subsequent motor fault vibration signal coding was carried out using GADF. The length of the data set to be encoded is usually 2^n, such as 64, 128, 256 and 512. As shown in Figure. 6, the accuracy rate of encoding using GADF under different data lengths is the highest when the data length reaches 256. When the data length exceeds 256, the accuracy rate will decline. At this time, each pixel in the feature map generated by encoding is compressed, which cannot reflect the characteristics of the original data very well, Therefore, the data length in the following text is 256.

Currently, most of the convolutional neural network structures use the Softmax function as the final classifier, but its sole purpose is to transform the results of the final fully connected layer into a probability distribution, which does not further improve accuracy. Here, ELM was chosen to replace the original classifier, making the diagnosis of the fault faster
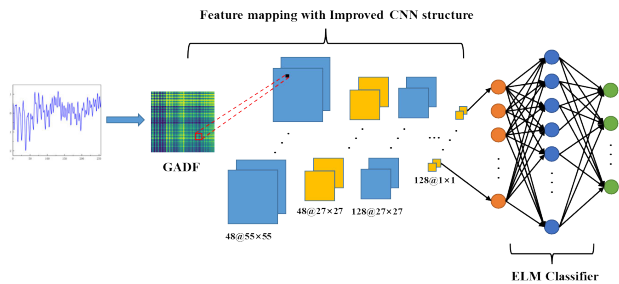
FIGURE 7. Architecture of the GCE.

FIGURE 8. The experimental platform used to produce the CWRU dataset [25].

and more accurate. The final structure of the GCE model can be seen in Figure.7.

The following is the overall flow of the GCE method:

Step 1: Collect the time series vibration signals of different faulty motors using accelerometric sensors.

Step 2: Use the Gram's difference angular field method to convert the 1-D time-series signal into a 2-D picture, preserving its original features to the maximum extent.

Step 3: Determine the CNN structure and network model parameters, and randomly initialize the weights of the CNN before starting training.

Step 4: Use ELM as a classifier to construct the complete GCE method and identify the number of nodes in ELM's hidden layer.

Step 5: In the proposed GCE, the images are initially fed into the CNN for training, and the network weight values of the best test results are retained. Since the ELM has the ability of fast convergence and the proposed CNN has retained the network weights of the best results, only a small amount of image input to the CNN for feature extraction is required to meet the ELM training requirements. In this paper, 10% of the images in the original sample are used as training inputs, and the number of ELM hidden layer nodes is adjusted in this process to determine the final parameters.

Step 6: The test dataset is fed into the trained GCE model to generate the final diagnostic results in the final testing phase.

IV. EXPERIMENTAL VERIFICATION

Two different motor bearing fault datasets were chosen to test the proposed method. The bearing datasets are from Case Western Reserve University (CWRU) [25]and autonomous experiment.

A. CWRU DATASET

1) DATASET DESCRIPTION

The CWRU dataset is a commonly used fault diagnosis dataset, and it was selected for comparative purposes. The experimental bench setup of the CWRU dataset which consists of motor, encoder, torque sensor, dynamometer and control electronics (Figure. 8). Its drive end and fan end use two different types of rolling bearings. In this paper, only the bearing failure signal collected from its drive end is used here, whose bearing type is 6205-2RS SKF. The bearing failures
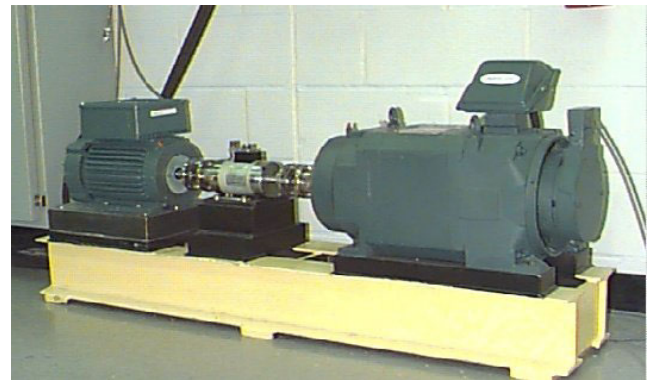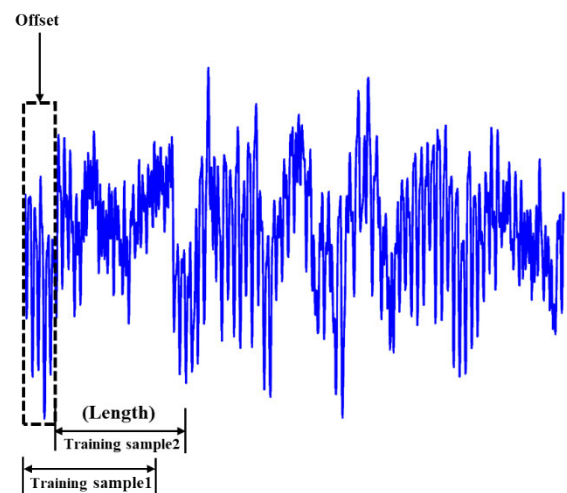

**FIGURE 9.** Dataset enhancement method.

are all made by electro-discharge machining (EDM), and the sampling frequency of the experimental platform is set to 12kHz. There are three types of motor bearing failure; rolling body damage, outer ring damage and inner ring damage, each damage diameter has three distinct sizes, which are 0.007 inches, 0.014 inches and 0.021 inches. They produce nine different failure states in total. The original experimental platform provides four different load conditions, and in this experiment, three of them are taken to create the dataset (1, 2 and 3hp), which corresponds to motor speeds of 1772, 1750, and 1730 rpm.

In this experiment, a total of three datasets A, B and C are prepared. As Table 3 shows, each dataset includes 18,000 training samples and 2,000 test samples, and the samples are all produced with the data enhancement method of overlapping sampling, with each acquisition length of 256 and an offset of 50, as shown in Figure. 9, and the acquired data are converted into 2-D pictures by the image coding method of GADF.

As shown in Figure. 10 for various fault types encoded by GADF images, there are 10 different states (including the normal state) under the same load, which correspond to each
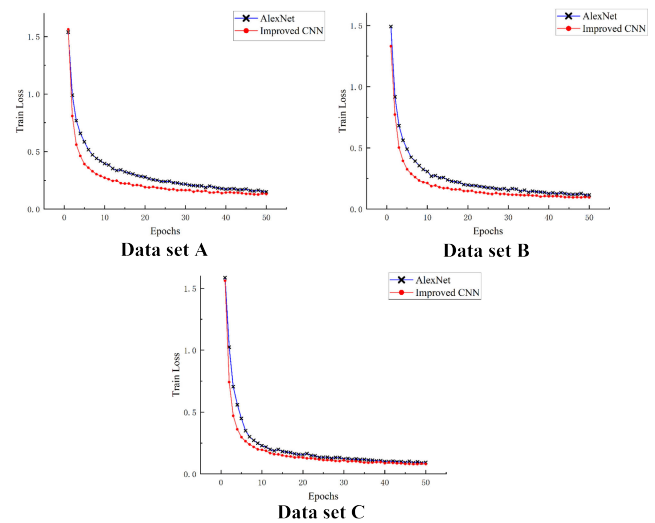
TABLE 3. Description of the CWRU dataset.

Fault type		Normal	Ball			Inner race			Outer race			Load
Label		1	2	3	4	5	6	7	8	9	10	
Fault diameter (inch)		0	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021	
A	Train	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	1hp
	Test	200	200	200	200	200	200	200	200	200	200	
B	Train	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	2hp
	Test	200	200	200	200	200	200	200	200	200	200	
C	Train	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	3hp
	Test	200	200	200	200	200	200	200	200	200	200	
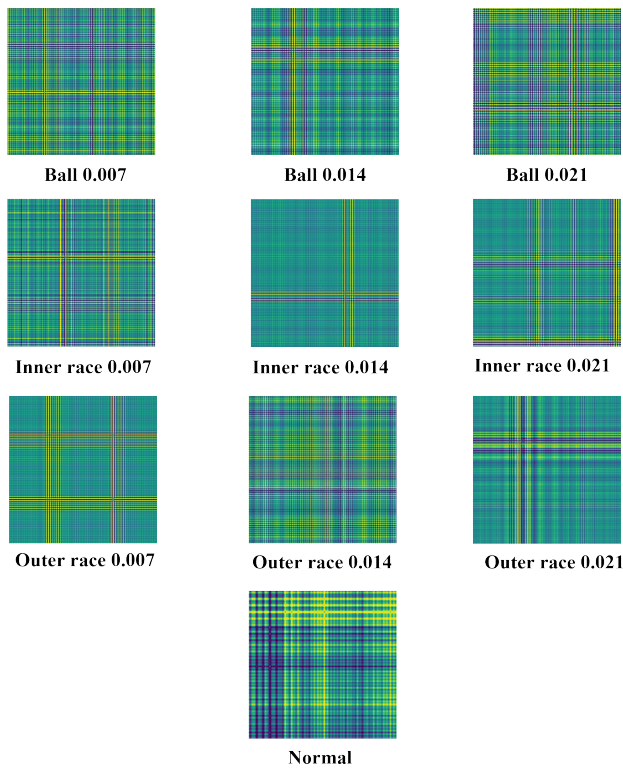
FIGURE 10. Fault characteristics of various types of bearing faults coded by GADF on the CWRU dataset.

FIGURE 11. Comparison of model training losses using the CWRU dataset.

other with the labels shown in Table 3. The Adam optimizer was chosen to be the model optimizer during training, and its learning rate was set to 0.001, the loss function was selected as the cross-entropy loss function, there were 18000 training images, the batch_size was set to 64, and the activation function was selected as the ReLu activation function. All calculations were completed on a personal computer with a CPU (AMD Ryzen 7 5800H@3.2GHz) and GPU (NVIDIA GeForce RTX 3060 6G).

2) VERIFICATION OF THE PROPOSED CONVOLUTIONAL NEURAL NETWORK'S EFFECTIVENESS

The three datasets A, B and C are trained with the AlexNet network model and modified convolutional neural network respectively. According to Figure. 11, the value of the loss function of the enhanced CNN model is lower than that of the

original AlexNet model at each iteration for all three datasets. The difference between the predicted value and the actual value is calculated using the loss function, and a lower value shows that the output of the model is nearer to the actual outcome, thus showing that the improved CNN model can converge faster and has better prediction performance than the AlexNet model.

To further evaluate the performance of the model, accuracy, precision, recall and F_1-measure are introduced as evaluation indexes, and the indexes are formulated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (15)$$

$$Precision = \frac{TP}{TP + FP} \quad (16)$$

$$Recall = \frac{TP}{TP + FN} \quad (17)$$

$$F_1 - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (18)$$

where TP represents a positive sample with correct model prediction, FP represents a positive sample with incorrect model prediction, TN represents a negative sample with correct model prediction, and FN represents a negative sample with incorrect model prediction.

TABLE 4. Model evaluation indexes using the CWRU dataset.

Data set	A		B		C	
	AlexNet	Improved CNN	AlexNet	Improved CNN	AlexNet	Improved CNN
Accuracy	0.9580	0.9750	0.9630	0.9785	0.9835	0.9850
Precision	0.9594	0.9759	0.9644	0.9799	0.9838	0.9851
Recall	0.9580	0.9750	0.9629	0.9785	0.9834	0.9850
F1-measure	0.9577	0.9747	0.9630	0.9783	0.9834	0.9850

TABLE 5. Diagnostic results of using different classifiers on the CWRU dataset.

Classifier	Testing accuracy (%)			Testing time (s)		
	A	B	C	A	B	C
Softmax	97.50	97.85	98.50	2.136	2.246	2.193
SVM	98.05	98.45	98.70	0.454	0.460	0.442
ELM	98.75	98.85	99.20	0.809	0.801	0.835
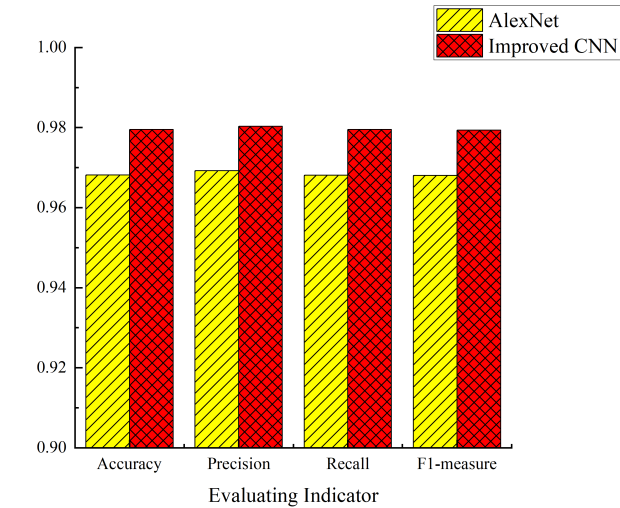
FIGURE 12. Comparison of the mean values of the indicators using the CWRU dataset.
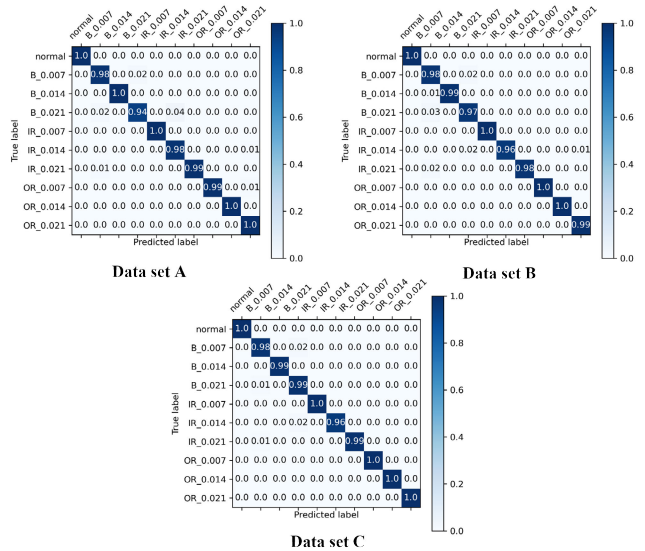
FIGURE 13. The GCE classification confusion matrix on the CWRU dataset.

To enhance the reliability of the evaluation metrics, 10 tests were conducted on each dataset for both AlexNet and the improved CNN model, and the final results of each index were averaged over 10 tests (Table 4).

For observation, histograms were drawn by fusing the indicators on the three data sets together and taking the average. As illustrated in Figure. 12, it can be intuitively found that the proposed CNN model is better than the original AlexNet model in all indexes, and with all other parameters being the same, the improved CNN model completes training 1.9% faster compared to the AlexNet model, with all other parameters being the same and with training on 18,000 samples. For the test with 2000 samples, the AlexNet model took 9.78s, while the improved CNN model took 9.23s, which saves 5.6% time. All of the above show the superiority of the improved CNN model, which can converge faster than the original model, and has improved model evaluation indexes.

3) VALIDATION OF THE EFFECT OF THE ELM CLASSIFIER

In deep learning fault diagnosis, the Softmax classifier is often used, although it only performs the transformation of probability distribution and requires multiple gradient updates to provide good classification results. The ELM however, can complete fault classification more quickly and accurately using Moore-Penrose generalized inverse matrix optimization. The number of neurons in the hidden layer of

the ELM is set to 10000. To assess the efficacy of ELM, the faulty vibration signal after completion of the GADF image coding is first input to the improved CNN model to extract features. The extracted advanced features are then fed into Softmax, SVM, and ELM classifiers for diagnosis.

Based on the results presented in Table 5, the test accuracy and test time of different classifiers on the three data sets A, B, and C from CWRU. To avoid randomness, all results are the average of 10 tests. Using SVM and ELM as classifiers can improve the diagnostic speed and diagnostic accuracy, and using ELM as the classifier has the highest accuracy, although the time is slightly increased compared to SVM. After comprehensive consideration, the best classifier that balances diagnostic accuracy and speed is ELM.

To further illustrate how well the improved approach works at spotting different kinds of faults, a confusion matrix was introduced to present the results, where each row of the confusion matrix represents the true attribution class for the data, and each column of the confusion matrix represents the predicted category. The confusion matrix of the predicted results of GCE on the three CWRU datasets is shown in Figure. 13. The dark part represents the accuracy of each type of fault being classified, and the rest of the values represent the percentage misclassified as other types. For example, on dataset A, 98% of the rolling body faults of 0.007 inches are correctly identified by the proposed method, and 2% are misclassified as rolling body faults of 0.021 inches.
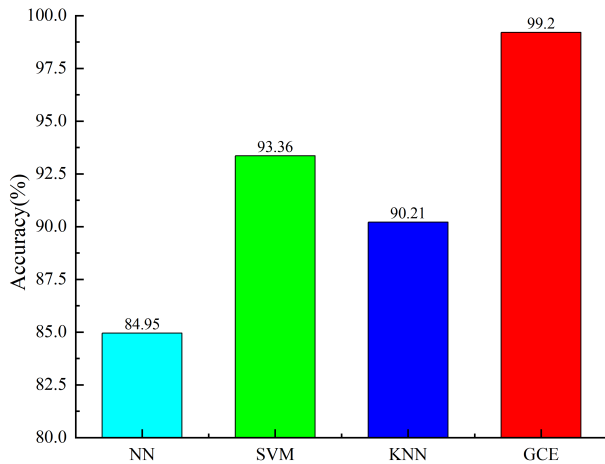
FIGURE 14. Diagnostic accuracy of different methods using the CWRU dataset.

FIGURE 15. The experiment platform of autonomous experimental dataset.

As illustrated in Figure. 13, the proposed method has good diagnostic capacity for the different fault types.

To further characterize the performance of the GCE method, three traditional intelligent diagnosis methods were used with the same data set for comparison; neural networks, support vector machines, and KNN (Figure. 4). Among the four algorithms, the proposed GCE approach has superior diagnostic accuracy.
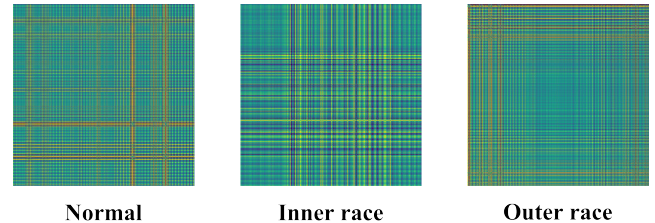
B. AUTONOMOUS EXPERIMENTAL DATASET

1) DATASET DESCRIPTION

For a further assessment of the performance of the proposed method, it was also validated by means of a motor rolling bearing failure dataset produced by autonomous experiments. The experimental bench consists of a detachable bearing housing, a sliding shaft, a detachable rotor with end circlips, a coupling, a pulley, a multi-stage belt drive-spur gearbox mechanism and a reciprocating mechanism. (Figure. 15).

This dataset uses rolling bearing type 6205. The motor bearing states are classified as normal, outer ring failure and

TABLE 6. Description of experimental data of the autonomous experimental dataset.

Fault type		Normal	Inner race	Outer race	Load
Label		0	1	2	
A'	Train	1800	1800	1800	1HP
	Test	200	200	200	
B'	Train	1800	1800	1800	2HP
	Test	200	200	200	
C'	Train	1800	1800	1800	3HP
	Test	200	200	200	

Normal   Inner race   Outer race

FIGURE 16. Fault characteristics of numerous bearing faults types coded by GADF on the autonomous experimental dataset.

Data set A'
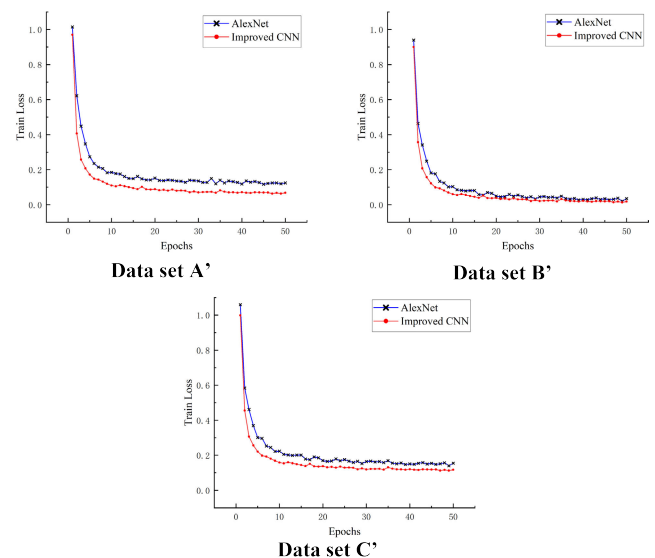
Data set B'

Data set C'

FIGURE 17. Comparison of model training losses using the autonomous experimental dataset.

inner ring failure. Data is collected by replacing the motor with different fault types at a sampling frequency of 12 kHz. 3 load states (1,2 and 3HP) are set up to create the data set for this experiment, corresponding to motor speeds of 1500, 1530 and 1570 rpm.

The number of neurons in the hidden layer of the ELM is set to 5000. This experiment also provided 3 datasets A', B', and C', as listed in Table 6, every dataset includes 5400 training samples and 600 test samples, and the samples are also produced with the same data enhancement method of overlapping sampling, and with the same acquisition length of 256 and offset of 50.(Figure. 9)

TABLE 7. Model evaluation indexes using the autonomous experimental dataset.

Data set	A'		B'		C'	
	AlexNet	Improved CNN	AlexNet	Improved CNN	AlexNet	Improved CNN
Accuracy	0.9725	0.9850	0.9885	0.9900	0.9817	0.9883
Precision	0.9733	0.9854	0.9889	0.9902	0.9820	0.9887
Recall	0.9724	0.9850	0.9886	0.9899	0.9817	0.9883
F1-measure	0.9721	0.9851	0.9887	0.9901	0.9817	0.9884

TABLE 8. Diagnostic results of using different classifiers on the autonomous experimental dataset.

Classifier	Testing accuracy (%)			Testing time (s)		
	A'	B'	C'	A'	B'	C'
Softmax	98.50	98.85	98.83	2.34	2.28	2.32
SVM	98.75	99.12	99.20	0.164	0.158	0.172
ELM	99.00	100.00	99.36	0.228	0.216	0.212
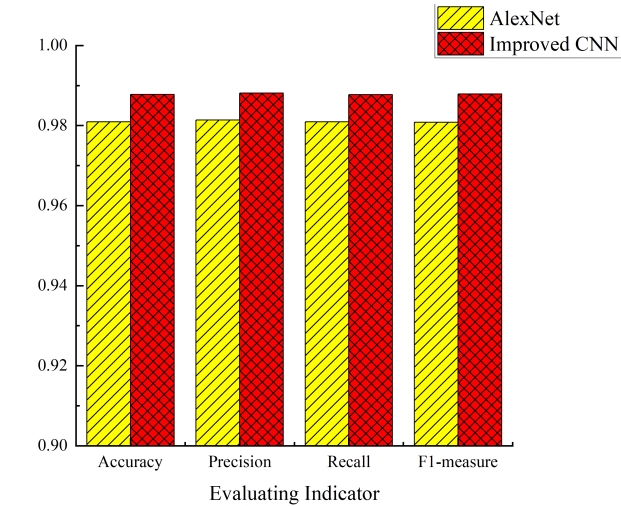
FIGURE 18. Comparison of the mean values of the indicators using the autonomous experimental dataset.
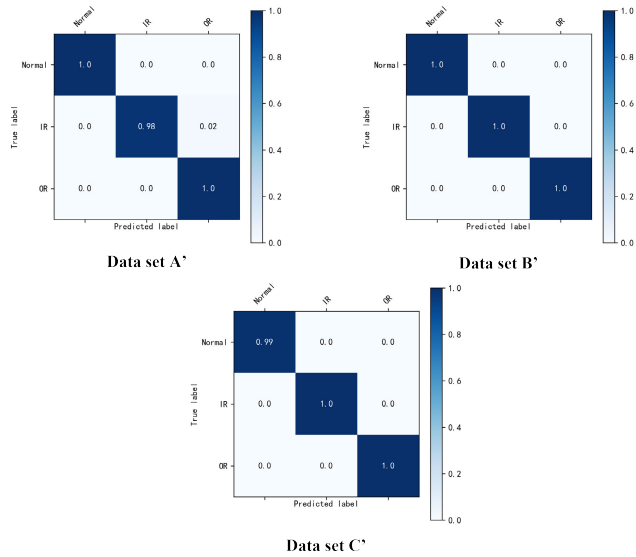
FIGURE 19. Classification confusion matrix of GCE using the autonomous experimental dataset.

The sample of each type of data after encoding by GADF is shown in Figure. 16. The optimizer, loss function, activation function selection and learning rate remain unchanged and there are 5400 training images, which are processed in small batches with batch size set to 32. The calculation environment is kept the same, which means the same personal computer.

2) VALIDATION OF THE EFFECT OF THE PROPOSED CONVOLUTIONAL NEURAL NETWORK STRUCTURE

AlexNet and the improved CNN network model were trained respectively by using the three datasets. As displayed in Figure. 17, the loss value of the improved CNN model proposed was still lower than that of the AlexNet model in each iteration, which once again verified the superiority of the proposed model.

Accuracy, precision, recall and F_1-measure were introduced as evaluation indexes, and each index was the mean value of each dataset after 10 tests, and the specific values are given in Table 7.

For the sake of observation, the histogram is still chosen to plot the mean value of each indicator on the three datasets. As shown in Figure. 18, the improved CNN model still outperforms the AlexNet model in all indexes, and the training time of the proposed CNN model is 5.2% faster than that of the AlexNet model for 5400 samples without changing other
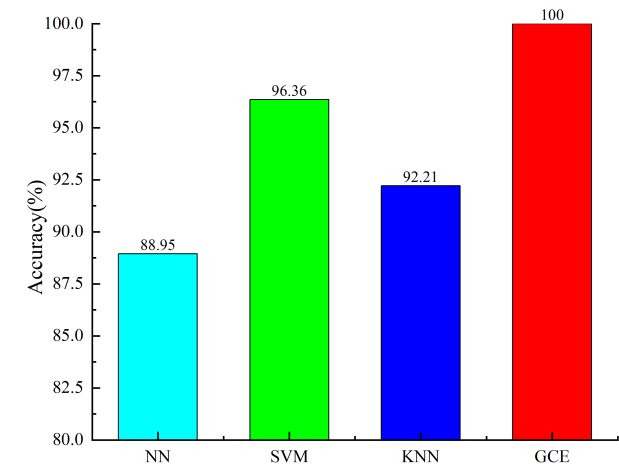

FIGURE 20. Diagnostic accuracy of different methods using the autonomous experimental dataset.

parameters. For a test of 600 samples, the AlexNet model takes 4.11s and the improved CNN model takes 2.32s, which is a 44% reduction in time required. In summary, all of the above shows that the improved CNN model is effective.

3) VALIDATION OF THE EFFECT OF THE ELM CLASSIFIER

Test accuracy and test time of different classifiers on three autonomous experimental datasets, designated A', B', and C',

are listed in Table 8. To avoid randomness, all results were averaged over 10 tests, and the best classifier that takes into account both speed and diagnostic accuracy is ELM.

Similarly, the confusion matrix was used to observe the classification effect of the proposed fault diagnosis method on various types of faults. The confusion matrices of the predicted results of GCE on the three datasets A', B' and C' from autonomous experiments, are depicted in Figure. 19. It is clear that the various fault types can be correctly categorized.

In addition, to describe the performance of the GCE approach in comparison to conventional methods, three traditional intelligent diagnosis methods were used for comparison with the same data illustrating the improved accuracy of the proposed method (Figure. 20).

V. CONCLUSION

This study proposes a new technique that combines GCE approaches to enable quick and precise motor bearing fault diagnosis. The first stage of the method involves GAF image coding on the vibration signal. To demonstrate that the suggested coding method is effective, time-frequency and grayscale maps are also introduced for comparison, and the GADF is the current optimal coding method based on its accuracy. The second stage is to improve the original CNN's network structure by substituting the global average pooling layer for the fully connected layer and introducing the batch normalization layer in the convolutional layer and the maximum pooling layer. After being trained on two different datasets, the training loss curve, various index parameters and the training times are compared, indicating that the improved network structure not only reduces the training parameters substantially, but also has a faster training speed and stronger generalization ability. Finally, the original classifier is replaced with the ELM classifier, leading to significant improvements in diagnosis speed and accuracy. Finally, the confusion matrix shows that the GCE can achieve effective identification for each type of fault.

To demonstrate the superiority of GCE comparisons with several traditional intelligent diagnosis algorithms were made by testing with two validation datasets from CWRU and autonomous experiments, and the results indicate that the GCE significantly improves the accuracy of diagnosis. In different datasets, GCE can accurately classify different types of faults, which fully reflects the advantages of the proposed method with strong generalization ability. Although there are many fault diagnosis methods that can achieve an acceptable level of accuracy, the GCE can achieve comprehensive fault diagnosis performance considering accuracy, speed and generalization ability.

REFERENCES

[1] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Dec. 2005.

[2] Y. Yang, M. M. M. Haque, D. Bai, and W. Tang, "Fault diagnosis of electric motors using deep learning algorithms and its application: A review," *Energies*, vol. 14, no. 21, pp. 1–27, Oct. 2021.

[3] 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. 2017.

[4] V. Nguyen, J. Seshadrinath, D. Wang, S. Nadarajan, and V. Vaiyapuri, "Model-based diagnosis and RUL estimation of induction machines under interturn fault," *IEEE Trans. Ind. Appl.*, vol. 53, no. 3, pp. 2690–2701, May 2017.

[5] 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. 2017.

[6] O. AlShorman, M. Irfan, N. Saad, D. Zhen, N. Haider, A. Glowacz, and A. AlShorman, "A review of artificial intelligence methods for condition monitoring and fault diagnosis of rolling element bearings for induction motor," *Shock Vibrat.*, vol. 2020, pp. 1–20, Nov. 2020.

[7] 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.

[8] D.-S. Huang, *Systematic Theory of Neural Networks for Pattern Recognition*. Beijing, China: Publishing House of Electronic Industry of China, 1996.

[9] H. Li, Q. Zhang, X. Qin, and Y. Sun, "Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network," *J. Vib. Shock*, vol. 37, no. 19, pp. 124–131, Sep. 2018.

[10] R. Rubini and U. Meneghetti, "Application of the envelope and wavelet transform analyses for the diagnosis of incipient faults in ball bearings," *Mech. Syst. Signal Process.*, vol. 15, no. 2, pp. 287–302, 2001.

[11] N. E. Huang, "New method for nonlinear and nonstationary time series analysis: Empirical mode decomposition and Hilbert spectral analysis," *Proc. SPIE*, vol. 4056, pp. 197–209, Apr. 2000.

[12] J. Harmouche, C. Delpha, and D. Diallo, "Incipient fault detection and diagnosis based on Kullback–Leibler divergence using principal component analysis: Part II," *Signal Process.*, vol. 109, pp. 334–344, Apr. 2015.

[13] Y. Guo, X. Wu, J. Na, and R. F. Fung, "Incipient faults identification in gearbox by combining Kurtogram and independent component analysis," *Appl. Mech. Mater.*, vols. 764–765, pp. 309–313, May 2015.

[14] F. Li, G. Meng, L. Ye, and P. Chen, "Wavelet transform-based higher-order statistics for fault diagnosis in rolling element bearings," *J. Vib. Control*, vol. 14, no. 11, pp. 1691–1709, 2008.

[15] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[16] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, Inception-ResNet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, no. 1, 2017, pp. 4278–4284.

[17] Q. Rao and J. Frtunikj, "Deep learning for self-driving cars: Chances and challenges," in *Proc. 1st Int. Workshop Softw. Eng. AI Auto. Syst.*, May 2018, pp. 35–38.

[18] T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *IEEE Comput. Intell. Mag.*, vol. 13, no. 3, pp. 55–75, Aug. 2018.

[19] P. Tamilselvan and P. F. Wang, "Failure diagnosis using deep belief learning based health state classification," *Rel. Eng., Syst. Safety*, vol. 115, pp. 124–135, Jul. 2013.

[20] F. Example, "Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings," *Shock Vib.*, vol. 2017, pp. 1–17, Oct. 2017.

[21] W. Gong, H. Chen, Z. Zhang, M. Zhang, R. Wang, C. Guan, and Q. Wang, "A novel deep learning method for intelligent fault diagnosis of rotating machinery based on improved CNN-SVM and multichannel data fusion," *Sensors*, vol. 19, no. 7, p. 1693, 2019.

[22] T. Han and Y.-F. Li, "Out-of-distribution detection-assisted trustworthy machinery fault diagnosis approach with uncertainty-aware deep ensembles," *Rel. Eng. Syst. Saf.*, vol. 226, Oct. 2022, Art. no. 108648.

[23] Z. Chen, K. Gryllias, and W. Li, "Mechanical fault diagnosis using convolutional neural networks and extreme learning machine," *Mech. Syst. Signal Process.*, vol. 133, Nov. 2019, Art. no. 106272.

[24] B. Han, H. Zhang, M. Sun, and F. Wu, "A new bearing fault diagnosis method based on capsule network and Markov transition field/Gramian angular field," *Sensors*, vol. 21, no. 22, p. 7762, Nov. 2021.

[25] *Case Western Reserve University Bearing Data Center Website*. Accessed: Oct. 25, 2022. [Online]. Available: http://csegroups.case.edu/bearingdatacenter/home

[26] Z. Wang and T. Oates, "Imaging time-series to improve classification and imputation," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, Jun. 2015, pp. 3939–3945.

[27] T. Han, T. Zhou, Y. Xiang, and D. Jiang, "Cross-machine intelligent fault diagnosis of gearbox based on deep learning and parameter transfer," *Struct. Control Health Monitor.*, vol. 29, no. 3, Mar. 2022, Art. no. e2898.

[28] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, *arXiv:1312.4400*.

[29] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, vol. 37, Jul. 2015, pp. 448–456.

[30] G.-B. Huang, "What are extreme learning machines? Filling the gap between Frank Rosenblatt's dream and John von Neumann's puzzle," *Cognit. Comput.*, vol. 7, no. 3, pp. 263–278, 2015.

[31] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[32] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

HAIBO GAO received the M.Sc. and Ph.D. degrees in marine engineering from the Wuhan University of Technology, Wuhan, China, in April 2000 and December 2008, respectively. He was a Visiting Scholar at Newcastle University, U.K., in 2013. He is currently an Associate Professor with the School of Naval Architecture, Ocean and Energy Power Engineering, Wuhan University of Technology. His research interests include system modeling and simulation, performance analysis of marine electric propulsion systems, and fault diagnosis.
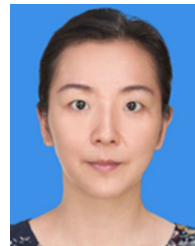
YELAN HE received the M.Sc. degree in computer application technology from Wuhan University of Technology, in April 2001. She is currently a Lecturer with the School of Computer and Artificial Intelligence, Wuhan University of Technology. Her research interests include virtual reality and fault diagnosis.

YUHAN ZHANG received the bachelor's degree in process equipment and control engineering from Jianghan University, China, in 2020. He is currently pursuing the master's degree in naval architecture with the School of Ship Ocean and Energy Power Engineering, Wuhan University of Technology, China. His research interests include power system fault diagnosis, pattern recognition, and artificial intelligence.

XIAOBIN XU received the Ph.D. degree in power electronics and power transmission from Shanghai Maritime University, China, in 2009. He is a Professor with the China–Austria Belt and Road Joint Laboratory on Artificial Intelligence and Advanced Manufacturing, Department of Automation, Hangzhou Dianzi University. His research interests include evidence theory, fuzzy set theory and applications in the processing of uncertain information, the reliability analysis, safety evaluation and condition monitoring of complex industrial systems.

LEI SHANG received the Ph.D. degree in marine engineering from the Wuhan University of Technology, Wuhan, China, in December 2003. She was a Visiting Scholar at the University of Maryland, USA, in 2011. She is currently a Professor with the School of Naval Architecture, Ocean and Energy Power Engineering, Wuhan University of Technology. Her research interests include system simulation and VR/AR technology.

YAJIE CHEN received the M.Sc. degree in marine engineering from the Wuhan University of Technology, China, in April 2000. She is currently a Researcher of marine engineering. Her research interests include new energy ship power systems, ship cabin automation, and intelligent cabin aspects.

• • •