RESEARCH PAPER

Convolution neural network model with improved pooling strategy and feature selection for weld defect recognition

Hongquan Jiang [1,2] · Qihang Hu [1] · Zelin Zhi [1,3] · Jianmin Gao [1] · Zhiyong Gao [1] · Rongxi Wang [1] · Shuai He [1] · Hua Li [1]

Abstract

Weld defect recognition plays an important role in the manufacturing process of large-scale equipment. Traditional methods generally include several serial steps, such as image preprocessing, region segmentation, feature extraction, and type recognition. The results of each step have significant impact on the accuracy of the final defect identification. The convolutional neural network (CNN) has strong pattern recognition ability, which can overcome the above problem. However, there are two problems: one is that the pooling strategy has poor dynamic adaptability and the other is the insufficient feature selection ability. To overcome these problems, we propose a CNN-based weld defect recognition method, which includes an improved pooling strategy and an enhanced feature selection method. According to the characteristics of the weld defect image, an improved pooling strategy that considers the distribution of the pooling region and feature map is introduced. Additionally, in order to enhance the feature selection ability of the CNN, an enhanced feature selection method integrating the ReliefF algorithm with the CNN is proposed. A case study is presented for demonstrating the proposed techniques. The results show that the proposed method has higher accuracy than the traditional CNN method, and establish that the proposed CNN-based method is successfully applied for weld defect recognition.

Keywords Weld defect recognition · Convolution neural network · Improved pooling strategy · Feature selection enhancement

Abbreviations

CNN	Convolutional neural networks
ReliefF	Feature selection algorithm
SVM	Support vector machine
CGP	Cartesian genetic programming
HD-CNN	Hierarchical Deep Convolutional Neural Network
P-CNN	Pose-based Convolutional Neural Network
RNN	Recurrent Neural Network
DCNN	Deep Convolutional Neural Network

✉ Hongquan Jiang
jhqxjtu@163.com

Qihang Hu
hqhexcellent@163.com

Zelin Zhi
zhi003@stu.xjtu.edu.cn

Jianmin Gao
gjm@mail.xjtu.edu.cn

Zhiyong Gao
zhygao@mail.xjtu.edu.cn

Rongxi Wang
rongxiwang@163.com

Shuai He
348580064@qq.com

Hua Li
1159822065@qq.com

1 State Key Laboratory for Manufacturing System Engineering, Xi'an Jiaotong University, Xi'an, China

2 Laboratory for Manufacturing and Productivity, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

3 Shaanxi Special Equipment Inspection and Testing Institute, Xi'an, China

ResNet	Residual network
RCNN	Region Convolutional Neural Network
ROI	Region of interest
F_{ij}	Input image
n	Moving step
S	Feature values obtained after pooling
σ_P	Value in the pooling region
σ_{FM}	Variance of the values on the feature map
t_{min}	Minimum value in the pooling region
t_{max}	Maximum value in the pooling region
t_{ave}	Average value considering the maximum and minimum values in the pooling region
μ	Based on the max pooling strategy
R	From the training sample set at a time
$W(A)$	Weight of each feature
m	Number of samples
$M_j(C)$	jth nearest neighboring sample in different categories of C
$p(C)$	Proportion of class C samples in the total
$\text{Class}(R_i)$	Category to which R_i belongs
$\text{diff}(A, R_i, R_j)$	Distance between R_i and R_j
C_n	Samples selected in category C
C_m	Total samples selected in all the categories
PO	Porosity
SL	Slag inclusion
LF	Lack of fusion
LP	Lack of penetration
CR	Crack

1 Introduction

In the manufacturing processes of major complex equipment, such as steam turbines and ships, welding is often involved; hence, weld defect analysis and recognition based on the ray detection image of the weld play an important role in guaranteeing the welding quality, reliability, and safety of the equipment [1, 2].

Currently, several studies based on traditional machine learning follow the traditional technical method, which mainly includes the process of "defect segmentation - feature extraction - feature selection - defect recognition" [3]. In these studies, modified background subtraction [4] and other methods are proposed for identifying the defect in the X-ray image; feature extraction generally obtains a set of edge-based features [5], hybrid descriptors based on the geometry [6], texture features [7], and other features. Feature selection primarily achieves the function of removing redundant features and noise to retain the useful features alone, and realizes the effective characterization of the defect-type features. Defect recognition is the effective judgment of the type and nature of the defect, based on the above steps, and is the core step in the entire defect recognition system, in which the Bayes [8], SVM [9], DS evidence theory [10], and other pattern recognition methods play an important role.

In recent years, with the development of artificial intelligence (AI) technology, methods based on the convolutional neural network (CNN) have become a research hotspot in image processing, pattern recognition, and other fields [11] because the end-to-end [12] recognition method addresses the issues involved in complex artificial processes, and have been applied in several fields such as environmental sound classification [13], grasp classification in myoelectric hands [14], and sentiment analysis [15]. In order to enhance the performance of the CNN, many improvements have been proposed; some of them focus on the common problems existing in the CNN. He et al. [16] proposed a pooling strategy called spatial pyramid pooling, for solving the problem of the artificial fixing of the size of the input image by the CNN, which decreases the recognition rate. Zhang [17] introduced a new graph CNN architecture based on the depth-based representation of a graph structure, which captures both the global topological structure and local connectivity structure within a graph. Suganuma [18] proposed a method for designing CNN architectures based on Cartesian genetic programming (CGP). Some studies have proposed CNNs for specific tasks: Yan et al. [19] proposed a HD-CNN for large-scale visual recognition, UberNet [20] for recognition tasks in computer vision, and P-CNN [21] for action recognition. HD-CNN reduces the top 1 and top 5 errors of VGG-19 model by 1.11% and 0.74%, respectively, and achieves advanced results on both CIFAR100 and large-scale ImageNet 1000-class benchmark datasets. In addition, in some studies, the improvement of the CNN is closely related to the characteristics of its application objects and the specific problems existing in the application of the CNN. Chaturvedi et al. [22] combined dynamic Gaussian Bayesian networks with the CNN to address an existing issue in previous works, wherein the prior distribution is not generally considered, when using sliding windows to learn word embedding. Ha et al. [23] introduced a multiple neural network topology, referred to as the selective deep CNN, to obtain accurate results for distorted images. Wang et al. [24] combined CNNs with the RNN and proposed a CNN-RNN framework to address the failure of explicitly exploiting the label dependencies in an image. The mean average precision of this method on PASCAL VOC 2007 dataset is 84%. Wu et al. [25] introduced a light CNN for deep face representation with noisy labels. Md Zahangir et al. [26] introduced an inception recurrent residual convolutional neural network (IRRCNN) model which is a deep convolutional neural network (DCNN) model. IRRCNN model combines the strengths of the inception network (Inception-v4), residual network (ResNet), and recurrent convolutional neural network (RCNN) for breast cancer classification. On the CIFAR-100 dataset, the IRRCNN model achieves 72.78% object recognition accuracy, which is about 4.53% higher than the recursive

convolutional neural network (RCNN). Abdulnabi et al. [27] proposed a joint multitask learning algorithm to better predict the attributes in images using deep CNN. The accuracy of attribute prediction of the algorithm proposed by the author is higher than other methods in color attributes, pattern group, cloth parts, and appearance group on the clothing dataset. Its total prediction accuracy reaches 92.82%. Chen et al. [28] presented an algorithm for unconstrained face verification, based on the CNN, for improving the performances of previous algorithms, which were often considerably degraded when images involving large variations in the pose, illumination, expression, aging, cosmetics, and occlusion were used. Recently, some studies have proposed convolution auditing neural networks for weld defect recognition. Khumaidi et al. [29] trained the CNN by replacing the convolution kernel with a Gauss kernel to form a neural network model to recognize two common types of weld defects. Liu et al. [30] proposed a VGG16-based fully convolutional structure for classifying weld defect images, achieving high accuracy with a relatively small dataset for the deep learning method. Anil et al. [31] improved the cost function of the CNN, which avoids redundant activation of the hidden layer in the CNN. Yuan et al. [32] changed the construction of the low and middle convolution kernels, and improved the generalization and convergence of CNNs. Xie et al. [33] combined data enhancement and the window slip detection method to realize defect classification and defect location marking. Using the super pixel segmentation algorithm and an improved ELU activation function, the model proposed by Fan et al. [34] can effectively identify four types of weld flaw detection images. And the overall recognition rate can reach 97.8%, which enhanced the recognition accuracy. Rui et al. [35] combined continuous wavelet transform (CWT) with the CNN to improve the accuracy. The accuracy of the proposed method by them is 96.94%, which is nearly 10% higher than the traditional method. Li et al. [36] constructed a deep learning network structure based on the principle of simulated visual perception, which can automatically learn the complex depth features in X-ray weld defect images.

However, in the existing studies on the application of CNNs for weld defect recognition, as the characteristics of the weld defect image are not studied in detail, the existing methods often lack pertinence and are therefore not conducive for the further improvement of the final defect recognition rate. First, the traditional pooling strategy (max pooling and average pooling) shows poor dynamic adaptability in the presence of different feature distributions in the weld defect area, resulting in inaccurate feature extraction for the entire image. As the gray distribution of the defects in a weld seam image has an important relationship with the gray distribution of the surrounding area [7], the pooling strategy needs to consider the distribution characteristics of the defect area. Furthermore, CNN-based methods generally include three layers: the input

layer, hidden layer, and output layer. The output feature vector of the hidden layer is an important factor that causes overfitting in the CNN model [11]. Therefore, improving the training ability and type recognition accuracy of the model by improving the feature selection ability of the output layer remains a problem to be solved.

To overcome the above problems, a CNN with improved pooling strategy, feature selection model, and weld defect recognition is proposed in this study. First, an improved pooling strategy considering the feature distribution of the pooling area and feature map comprehensively is proposed, which can overcome the problem in the traditional pooling strategy wherein the weld defects characteristics are disregarded. Furthermore, the ReliefF algorithm is integrated with the CNN for constructing a strengthened feature selection method. A CNN is then constructed and trained with the above pooling strategy and feature selection method for image recognition. A practical case demonstrates that this method effectively overcomes the shortcomings of the traditional CNN, improves the accuracy of the pool feature selection and feature selection ability of the CNN model, and achieves good recognition accuracy.

The remainder of the paper is organized as follows. Section 2 analyzes the two problems in the application of the CNN for weld defect recognition. Section 3 describes the proposed pooling strategy for solving the problem in the tradition pooling strategy, discusses the enhanced feature selection method, and illustrates the flow of the weld defect recognition method proposed in this study. Section 4 presents a weld defect recognition case and the results of the proposed method in comparison with the traditional CNN methods. Finally, Section 5 presents the conclusions on the proposed method.

2 Analysis of the problem in weld defect identification using a CNN

In this section, the problems in the pooling strategy and feature selection of the traditional CNN model are analyzed. The CNN, first proposed by Fukushima in 1980 [37], contains an input layer, hidden layer, and output layer. The hidden layer is generally composed of multiple convolution layers and polling layer structures, and a fully connected layer. The polling layer is obtained by pooling the input feature map. Selecting different continuous ranges in the input feature map as the pooling region, an $n \times n$ rectangular area is generally selected, and a feature in the pooling region is selected as the characteristic of the pooling region, for a certain strategy. The traditional feature selection strategy includes max and average pooling: max pooling involves the selection of the maximum value in the pooling region as the feature of the pooling region, whereas average pooling involves the

selection of the average value of the pooling region as the feature of the pooling region. Assuming that F_{ij} is the input image, $n \times n$ is the size of the pooling region, and n is the moving step, S represents the feature values obtained after pooling. The max and average pooling strategies are depicted in Eqs. (1) and (2), respectively.

$$S = \max_{i=1, j=1}^{n} \left(F_{ij} \right) \qquad (1)$$

$$S = \frac{1}{n^2} \left(\sum_{i=1}^{n} \sum_{j=1}^{n} F_{ij} \right) \qquad (2)$$

To demonstrate the problems involved in the feature extraction of the two traditional pooling strategies, two different representative pooling regions were selected, as shown in Fig. 1. Figure 1(a) shows the pooling process of the CNN model, in which the gray background area in the red box represents a 2 × 2-sized pooling region. Figure 1 (b) and 1(c) display two types of defects with the inclusion of slag and tungsten, respectively. The red and blue boxes in the images represent the pixel value distribution in the four pooling regions depicted by Fig. 1(d), 1(e), 1(f), and 1(g), respectively. Based on the distribution of the pixel values in Fig. 1(d)–1(g), two different pooling regions with different feature distributions can be observed because of the different positions of the pooling regions on the feature map. One is the pooling region depicted in Fig. 1(d) and 1(f). The distribution of the pixel values in this type of pooling region is more uniform. Pooling regions with such

feature distribution occur mostly in the weld area. Here, feature extraction by average pooling is appropriate. If maximum pooling is used, noise may be introduced. Another type of pooling region is depicted in Fig. 1(e) and 1(g), which is located in the edge zone of the defect edge. This pooling region contains the edge feature of the defect. If average pooling strategy is used to extract the features of this type of pooling region, it will result in the elimination of the edge features of the defect. Therefore, the traditional maximum pooling and average pooling strategies have poor dynamic adaptability for pooling regions with different feature distributions, resulting in inaccurate feature extraction. In a weld image with defects, although the variation of the gray values within and outside the defect area is different, the traditional maximum pooling and average pooling strategies do not reflect these variations.

Regarding the feature selection problem in the CNN, as previously mentioned, the output feature vector of the hidden layer has significant influence on the training ability and classification effect of the CNN model. Figure 2 shows the hidden layer structure of a typical CNN, in which W_{11}, W_{12}, W_{21}, and W_{22} constitute the parameter matrix of the convolution kernels on two layers of the convolution layer, and W_3 is the parameter transfer matrix between the two fully connected layers. In the CNN training process, the final output feature is determined and selected by continuously updating the parameters of W_{ij}. However, because CNN training often includes overfitting or under-fitting, the output feature includes redundancy and noise, reducing the efficiency of the final classification.

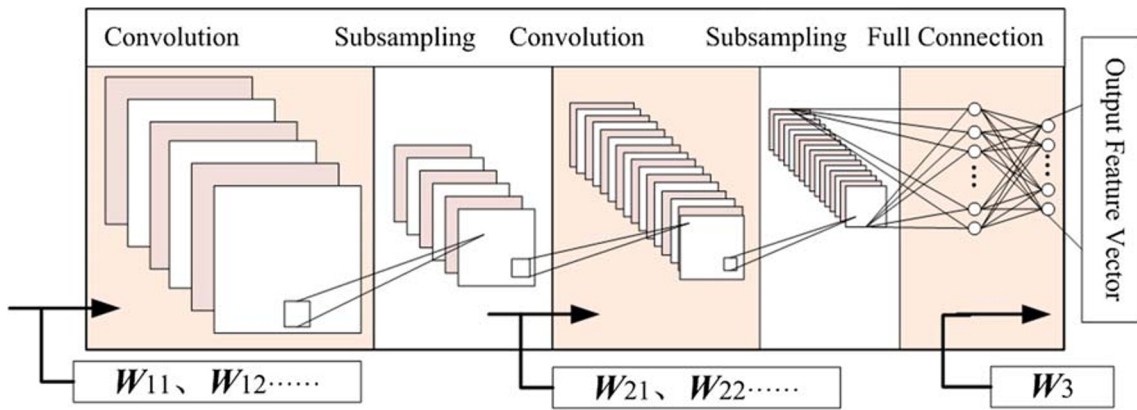**Fig. 1** Pooling strategies in different areas
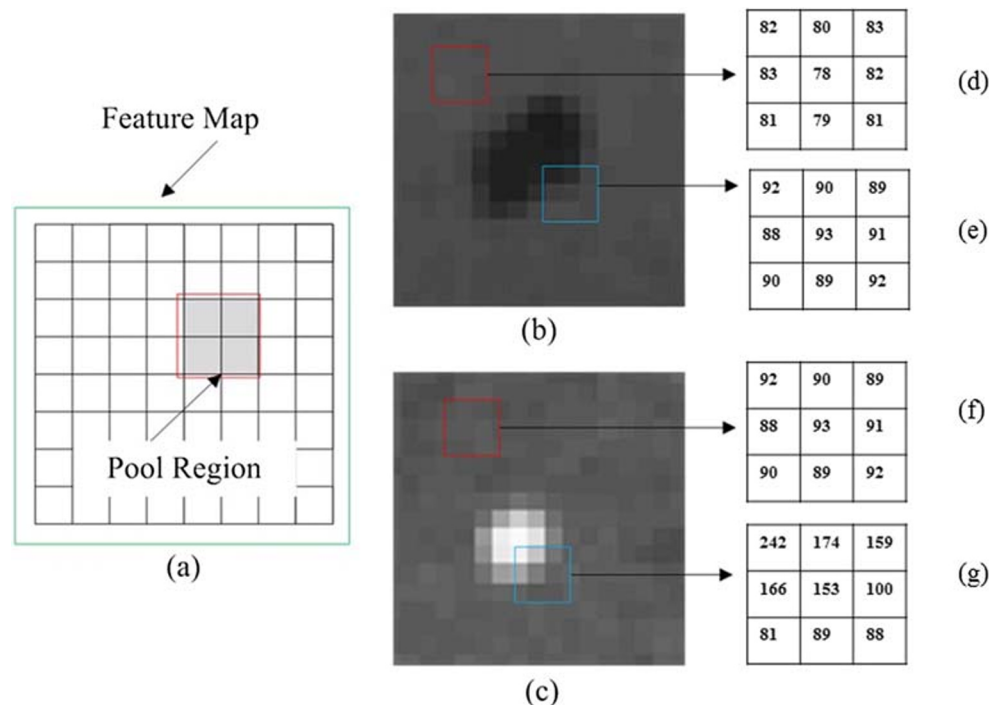
Fig. 2 Hidden layer structure of the CNN

3 Improved CNN model for weld defect–type identification

3.1 Improved pooling strategy

In this section, an improved pooling strategy is proposed to overcome the problems in the feature extraction of weld defect images using the traditional pooling strategy. The proposed improved pooling strategy is discussed using two types of defect images (slag inclusion and tungsten trapping) as examples, which are shown in Fig. 3; Fig. 3(a)–3(d) indicate the respective pooling domains located in four different regions of the two defect images. Different pooling methods are needed depending on whether the pooling domain is outside the defect area or on the edge of the defect area, and the calculation method is shown in Eq. (3).

$$S = \begin{cases} \min\limits_{i=1,j=1}^{n}\left(F_{ij}\right), \sigma_P \geq \sigma_{FM} \, and \, |t_{\min}-t_{ave}| > |t_{\max}-t_{ave}| \\ \max\limits_{i=1,j=1}^{n}\left(F_{ij}\right), \sigma_P \geq \sigma_{FM} \, and \, |t_{\min}-t_{ave}| \leq |t_{\max}-t_{ave}| \\ \mu \max\limits_{i=1,j=1}^{n}\left(F_{ij}\right), \sigma_P < \sigma_{FM} \end{cases} \quad (3)$$

where σ_P is the value in the pooling region, σ_{FM} is the variance of the values on the feature map, $t_{\min}$ is the minimum value in the pooling region, $t_{\max}$ is the maximum value in the pooling region, and t_{ave} is the average value considering the maximum and minimum values in the pooling region.

According to Eq. (3), when $\sigma_P \geq \sigma_{FM}$ and $|t_{\min} - t_{ave}| > |t_{\max} - t_{ave}|$, as in the case of Fig. 3(a), the minimum value in the pooling region is the output feature. When $\sigma_P \geq \sigma_{FM}$ and $|t_{\min} - t_{ave}| \leq |t_{\max} - t_{ave}|$, as in the case of Fig. 3(b), the maximum

Fig. 3 Revised max pooling strategy
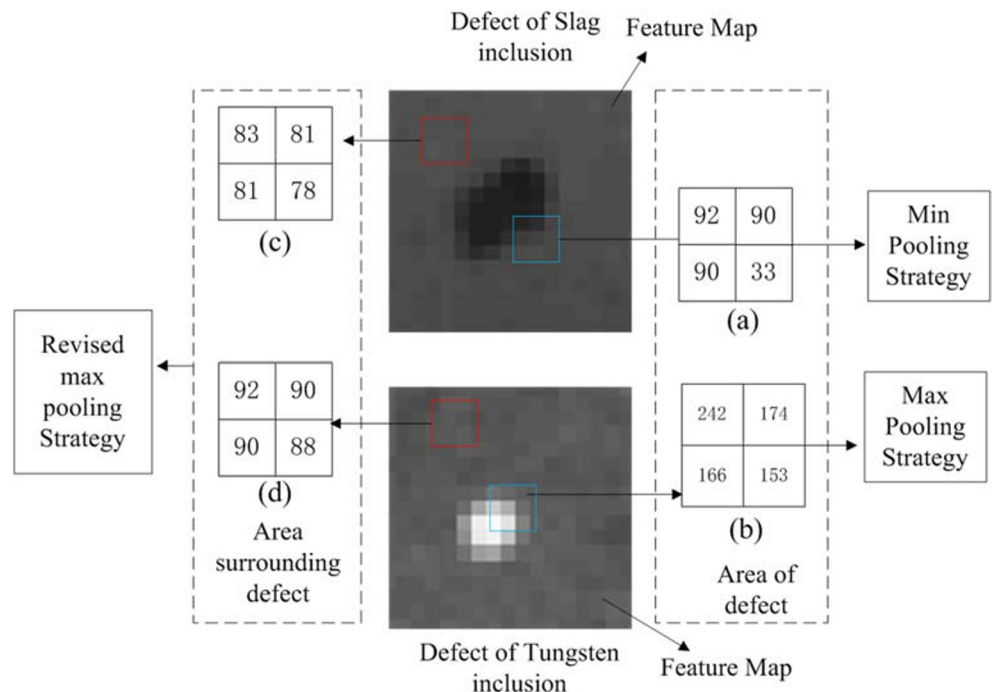
value in the pooling region is the output feature value. When $\sigma_P < \sigma_{FM}$, the feature variance in the pooling region is small and the maximum feature value is not obvious; for this case, a strategy is proposed, in which a modification factor, μ, is introduced based on the max pooling strategy. This modification factor is the ratio of the sum of the feature values and the difference in the pooling region to the sum of feature values in the pooling region, as shown in Eq. (4):

$$\mu = (t_{\mathrm{sum}} - t_{\max} + t_{\min})/t_{\mathrm{sum}} \tag{4}$$

For example, Fig. 3(a) shows the pooling strategy for the pooled region in the slag inclusion area, where the minimum feature value in the pooling region represents the defect edge characteristics (the feature values in the pooling region of Fig. 3(a) are 92, 90, 90, and 33, of which 33 represents the defect edge). Figure 3(b) shows the pooling strategy for the pooled region in the inclusion defect region, where the maximum feature values in the pooling domain indicate the edge characteristics of the tungsten inclusion defects (the feature values in the pooling domain of Fig. 3(b) are 242, 174, 166, and 153, of which 242 represents the feature values of the defect edge in the pooling domain). Therefore, the minimum and maximum pooling strategies can be adopted for Fig. 3(a) and 3(b), respectively, because their output feature values are 33 and 242, respectively. When the pooling region is located in the feature map area outside the defect area, as shown in Fig. 3(c) and 3(d), the output value is calculated using the improved maximum pooling strategy, namely, Eq. (3).

In summary, using the improved pooling strategy, different output feature value calculation methods are used according to the different locations of the pooling region in the feature map, which can reflect the characteristics of the defect image and include certain adaptability.

3.2 Enhanced feature selection method

To enhance the CNN feature selection ability, an enhanced feature selection method that integrates the ReliefF algorithm with the CNN is proposed. The ReliefF algorithm, first proposed by Gore [38], is a traditional feature evaluation method that deals with two classification problems. It can provide the corresponding weight, based on the significance of the feature. The greater the weight of the feature, the stronger is its classification ability. The ReliefF [39] algorithm can deal with multiclass problems, and is used for re-evaluating the features extracted by the CNN for recognition. When dealing with multiclass problems, the ReliefF algorithm randomly extracts a sample, R, from the training sample set at a time, and then finds the k nearest neighboring samples of R from the same sample set as R, and the k nearest misses from the different sample sets of each R, and then updates the weight of each feature, $W(A)$, as shown in Eq. (5):

$$W(A) = W(A) - \sum_{j=1}^{k} \mathrm{diff}(A, R, H_j)/\mathrm{mk}$$

$$+ \sum_{C \neq \mathrm{class}(R)} \left[\frac{p(C)}{1 - p(\mathrm{Class}(R_i))} \sum_{j=1}^{k} \mathrm{diff}(A, R, M_j(C)) \right]/(\mathrm{mk}) \tag{5}$$

where m is the number of sampling times, $M_j(C)$ is the jth nearest neighboring sample in different categories of C, $p(C)$ is the proportion of class C samples in the total, $\mathrm{Class}(R_i)$ is the category to which R_i belongs, and $\mathrm{diff}(A, R_i, R_j)$ is the distance between R_i and R_j, and is mathematically expressed by Eq. (6):

$$\mathrm{diff}(A, R_i, R_j)$$
$$= \begin{cases} \dfrac{\left| R_i[A] - R_j[A] \right|}{\max(A) - \min(A)}, & A \text{ is continuous} \\ 0, & A \text{ is discontinuous}, R_i[A] = R_j[A] \\ 1, & A \text{ is discontinuous}, R_i[A] \neq R_j[A] \end{cases} \tag{6}$$

The traditional ReliefF algorithm may cause the samples to fall into one or several categories, during the random sampling of multiclass samples, and the distribution of the characteristics of the entire sample cannot be considered. Based on this, this study adopts the "interclass ratio, intraclass randomness" sampling method; "interclass ratio" is the ratio of the number of samples extracted in each category to the total number of samples in that category:

$$\frac{C_n}{C_m} = \frac{n}{m} \tag{7}$$

where n is the total number of samples in category C, m is the total number of samples in all the categories, C_n represents the samples selected in category C, and C_m represents the total samples selected in all the categories; "intraclass randomness" refers to the random selection of samples in a category, whereas "intraclass randomness " refers to the random selection of samples in category C.

The feature weights calculated by ReliefF may contain negative values. If the feature weight is a negative value, it indicates that the distance between samples of the same category is greater than those between samples of different categories, which is contrary to the expected feature properties. Therefore, during feature selection, this study eliminates this type of feature and sets the corresponding weight of the feature to zero; the revised weight vector is then obtained, based on the initial weight vector provided by the ReliefF algorithm. The revised weight vector is assigned to the CNN for extracting the feature vectors for classification, and the features selected by combining the ReliefF algorithm with the CNN are obtained. Hence, the ReliefF algorithm combines its understanding of the feature significance with that of the CNN. The evaluation and selection of features are beneficial for improving the feature selection ability of CNN models.
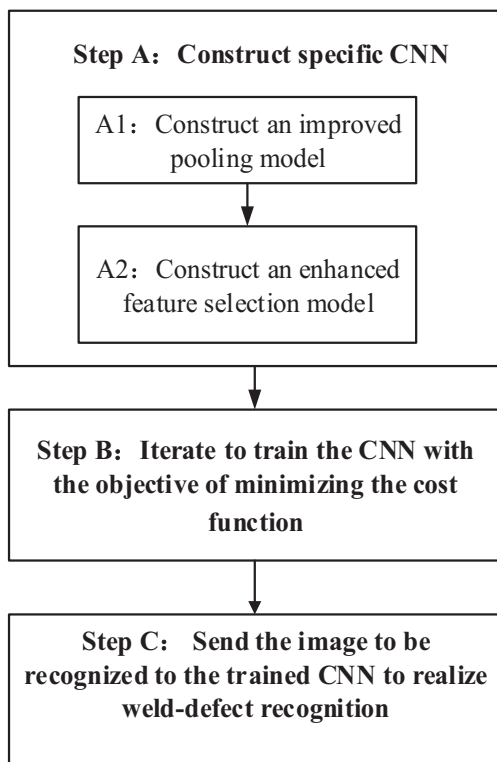
Fig. 4 Flow chart of the proposed defect recognition method

The basic structure of the network includes an input layer, two convolution layers (C1 and C3), two pooling layers (P2 and P4), two full connection layers (F5 and F6), one feature selection layer, and one output layer. By abstracting and extracting the input image layer-by-layer, the characteristic information of the representative sample can be obtained.

In the first convolution operation, the input image is convoluted by six convolution checkers sized 5 * 5, and six 28 * 28-pixel feature maps are obtained in the C1 layer. The input layer is a gray image of 32 * 32 pixels, with a stride of one and padding of zero. In the first pooling operation, six feature maps of the C1 layer are operated using a 2 * 2 pooling domain, and six 14 * 14-pixel feature maps are obtained in the P2 layer, with a stride of two.

In the second convolution operation, sixteen convolution checkers sized 5 * 5 are used to convolute the feature images of the P2 layer, and sixteen 10 * 10-pixel feature maps are obtained in the C3 layer, with a stride of one and padding of zero. In the second pooling operation, the feature maps of the C3 layer are operated using a 2 * 2 pooling domain, and sixteen 5 * 5-pixel feature maps are obtained in the P4 layer, with a stride of two.

The F5 and F6 layers are fully connected, containing only a one-dimensional vector. In the full connection operation of P4 and F5, 120 convolution checkers sized 5 * 5 are used to convolute sixteen feature maps of the P4 layer. The stride is one and the padding is zero. One hundred and twenty feature maps sized 1 * 1 are obtained in the F5 layer, i.e., the F5 layer is a one-dimensional vector containing 120 values. The F5 layer is fully connected, and a one-dimensional vector containing 84 values is obtained in the F6 layer.

The method described in Section 3.2 is used for feature selection of the F6 layer. Features with high importance are reserved, whereas those with low importance are eliminated by setting the corresponding node value to zero to get the feature selection layer. Finally, the softmax multiclass classifier is used to classify the feature selection layer, and the output layer is obtained. The principle of the softmax classifier is as follows:

$$P(Y = Y_i) = \frac{e^{Y_i}}{\sum_{i=1}^{K} e^{Y_i}},$$

3.3 Weld defect recognition process based on the proposed CNN

Based on the methods described in Sections 3 and 4, an improved CNN model for weld defect–type identification is proposed. The flow chart of the defect-type identification using this method is depicted in Fig. 4.

From the flow chart, we can see that the proposed defect recognition process includes steps A, B, and C. In step A, a CNN with a specific architecture is constructed, which includes substep A1 for constructing an improved pooling model considering the pooling region and area surrounding the defect feature distribution comprehensively, and substep A2 for constructing an enhanced feature selection method. The structure of the CNN with the improved pooling strategy and feature selection model is depicted in Fig. 5.

Fig. 5 Structure of the CNN with the improved pooling strategy and model

Fig. 6 X-ray film scanner (JD-RTD)

4 Experiment and result

The main research object of this study is the defect in the welding process of steam turbines, and radiographic images were provided by the Dongfang Turbine Co., Ltd., Sichuan, China. The base metal of the welding seam includes mainly steel, nickel, and copper, and the welding joint is a double-sided butt weld. As the weld defects are detected through X-ray inspection, the X-ray film is digitized using an X-ray film scanner (JD-RTD) developed in-house, as shown in Fig. 6.

There are five types of weld defects in digital radiograph, including porosity (PO), slag inclusion (SL), lack of penetration (LP), lack of fusion (LF), and crack(CR), as shown in Fig. 7.

Because the weld defect image was large, it was difficult to directly use the image as the input to the training network model of the neural network. Therefore, the original weld image was preprocessed, the defect and surrounding area in the original weld image was intercepted as a 32×32-sized region of interest (ROI), and the ROI image was the input to the neural network. In this study, 3486 ROI weld images were selected, including 504 porosity (PO), 410 slag inclusion (SL), 460 lack of fusion (LF), 864 lack of penetration (LP), 804 crack (CR), and 444 nondefect images. All these images were divided into a training set and testing set at a ratio of 4:1. A total of 2789 images were obtained for training and 697 images for testing. Some of the images, as experimental samples, are shown in Fig. 8.

The experiment was performed on a Windows 7 operating system using an Intel (R) Core (TM) i5-4460 CPU with a

where Y_i represents the output result corresponding to defect i in the image, $i = 1, 2, 3, 4, 5, 6$ corresponds to crack, lack of fusion, lack of penetration, slag inclusion, porosity, and nondefect in the weld, respectively. K is the number of categories and $P(Y=Y_i)$ is the probability information of the output results Y_i corresponding to defect i in the image.

Furthermore, in step B, iteration is performed with the objective of minimizing the cost function to train the neural network constructed in step A for weld defect recognition. Finally, in step C, the sample weld image to be identified is input to the CNN trained in step B for the automatic recognition of the defect types.

Fig. 7 Defects: (a) porosity (PO), (b) slag inclusion (SL), (c) lack of penetration (LP), (d) lack of fusion (LF), and (e) crack (CR)
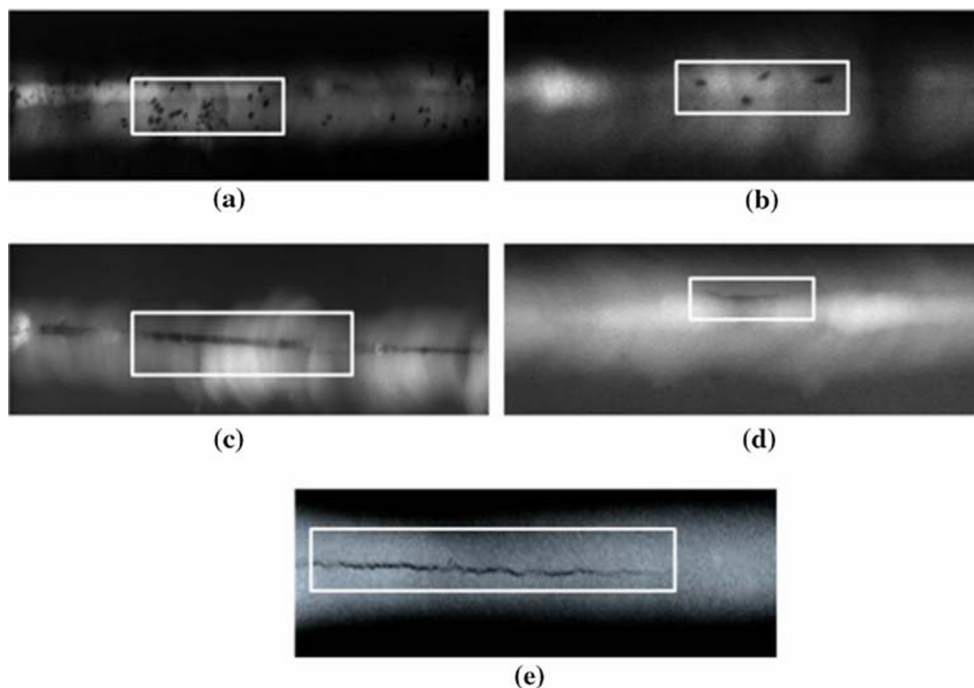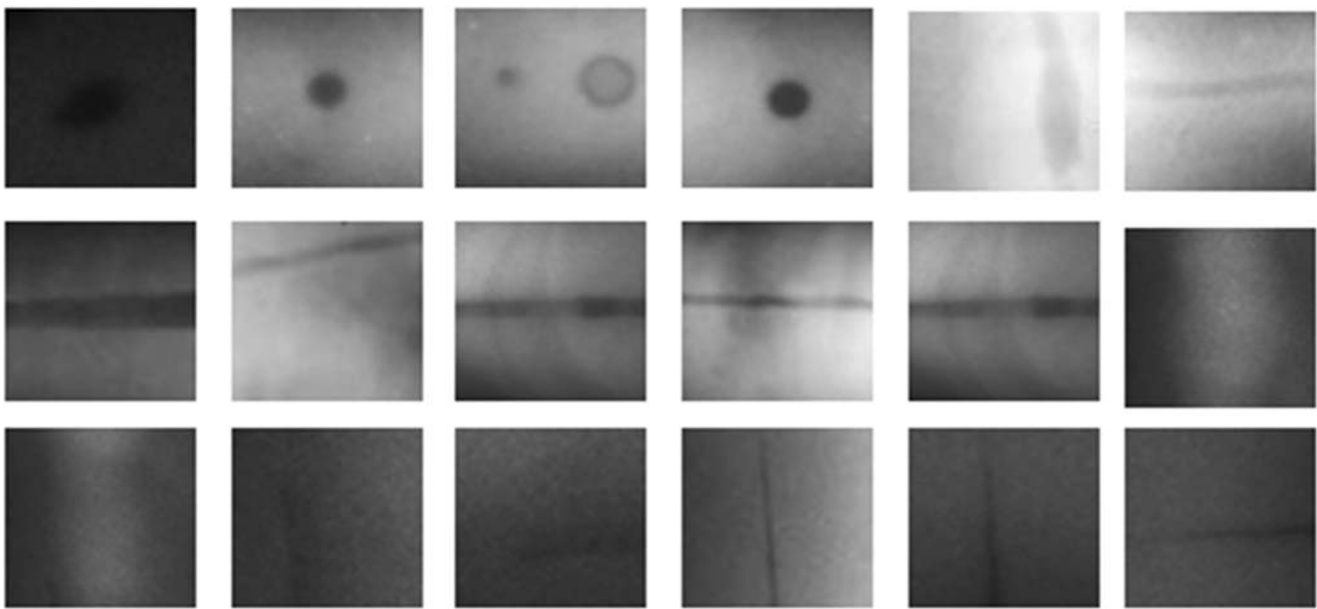
Fig. 8 Experimental sample images

3.20-GHz processor, 8.00-GB running memory, PyCharm integrated development environment based on Python 3.6.4, and Google open source tensorflow 1.13.0 deep learning framework. The model training and testing scheme is shown in Fig. 9.

The minibatch gradient descent method was selected for training; the batch size was set to 64 and cross entropy was used as the loss function. In the model training process, 20 steps were set as an epoch and the maximum number of training steps was 200. After training, the data of the test set was input to testing, and the accuracy of defect identification in the test set was obtained.

4.1 Validation of the improved pooling strategy

In order to verify the effectiveness of the pooling strategy proposed in this study, the basic CNN network model was used for the experiment, and the maximum pooling strategy, average pooling strategy, and pooling strategy proposed in this study were used to construct the network model in the pooling layer. And CNN-1, CNN-2, and CNN-3 were

obtained. The network architecture of the model is shown in Fig. 10. The selection of specific pooling strategies is shown in Table 1.

Experiments were performed using CNN-1, CNN-2, and CNN-3, respectively. The defect recognition accuracies, under different iterations, for the three models are listed in Table 2.

The experimental results demonstrate that under different iterations, pooling model, CNN-3, obtained higher accuracy than the max pooling model, CNN-1, and the average pooling model, CNN-2. The recognition accuracy calculation method is as follows: the single defect identification accuracy (correct number / total number) is calculated, and the average identification accuracy of all the defects is considered as the recognition accuracy.

According to Table 2, the accuracy of defect image recognition of the network models constructed using the pooling strategy is higher than those of the maximum pool strategy and average pool strategy models under different iterations. When the number of iterations is 200, the accuracy of defect image recognition of the network model constructed using the

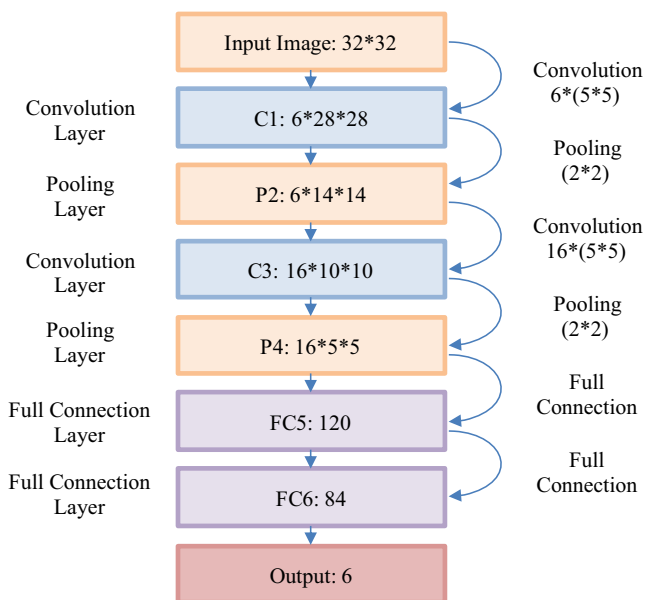**Fig. 9** Flow chart of the model training and testing scheme
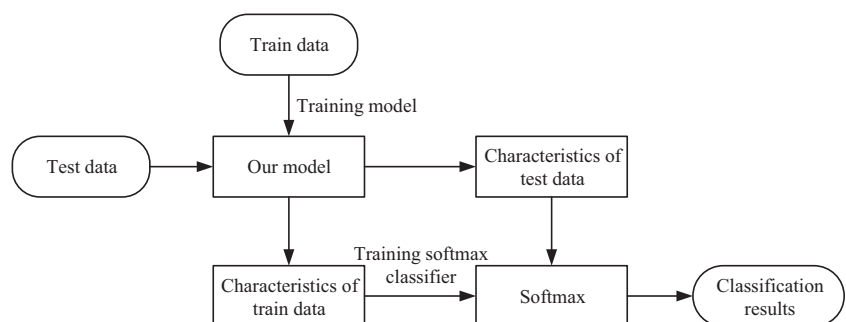
Fig. 10 Basic CNN architecture

pooling strategy is 90.0%, which is 5.3% higher than the model with the maximum pool strategy and 1.8% higher than the model with average pool strategy.

4.2 Validation of the enhanced feature selection method

To validate the proposed feature selection method, a feature selection layer was added between the FC6 layer and output layer of the basic CNN architecture (Fig. 10) to construct a CNN with enhanced feature selection (Fig. 11). In the feature selection layer, the features extracted from the FC6 layer were selected by the ReliefF algorithm. Features with strong classification ability were retained along with their weights, and the redundant features were eliminated. The selected features were used as the input to the output layer.

The training set included 3348 images. In the FC layer, 84 features were extracted from each image. In each iteration of the CNN shown in Fig. 11, the batch size used was 110, to obtain a feature set, T, of dimension, 110×84. The ReliefF algorithm described in Section 4 was then used for processing

Table 1 Pooling strategy selection

CNN architecture	Layers	
	Pooling strategy	
	P2	P4
CNN-1	Max pooling	Max pooling
CNN-2	Average pooling	Average pooling
CNN-3	Proposed strategy	Proposed strategy

Table 2 Recognition accuracies of the three pooling strategies under different iterations

CNN model	Times						
	Accuracy						
	50	75	100	125	150	175	200
CNN-1	0.828	0.846	0.846	0.864	0.855	0.864	0.855
CNN-2	0.848	0.857	0.864	0.864	0.875	0.875	0.884
CNN-3	0.866	0.866	0.875	0.884	0.893	0.900	0.900

feature set, T, and 84 feature weights were obtained. Some of the parameters in the ReliefF algorithm were set as follows: The number of neighboring samples in each sample set was five, and the number of samples selected randomly by the ReliefF algorithm for weight evaluation in each sample set is presented in Table 3.

After the above processing, we obtained the initial feature value vector, L_0, with 84 nodes corresponding to each feature of the FC6 layer in the CNN architecture shown in Fig. 11, with each iteration of the CNN; the 200th iteration is shown in Table 4.

From L_0, it was determined that the weights corresponding to some of the features were negative, indicating that the distances between samples in the class were greater than those between classes, which is not conducive for classification and needs to be eliminated. Based on this, feature weights less than zero were set to zero, and the revised weight vector, L_1, was obtained, as depicted in Table 5.
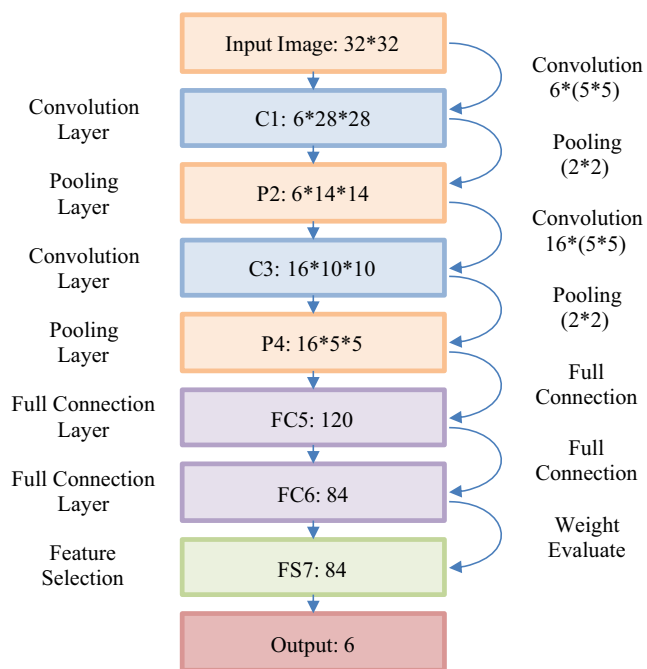

Fig. 11 Enhanced feature selection CNN architecture

Table 3 Number of extractions of different samples

Sample class	Total sample size	Number of times selected
Crack	643	26
Lack of fusion	368	14
Lack of penetration	691	28
Slag inclusion	328	12
Porosity	404	16
Nondefect	355	14
Total	2789	110

The feature weight of weight vector, L_1, was assigned to the corresponding feature in FC6 of the CNN architecture shown in Fig. 11, and the feature selected by the ReliefF algorithm could be obtained at the feature selection level. For the zero weights in L_1, the FC6 feature was given a corresponding feature weight, which is equivalent to removing the corresponding feature with a weight of zero. A total of 16 zero feature weight values were calculated, indicating that the feature vectors in the FC6 layer were reduced to 68 dimensions after feature selection.

As shown in Table 5, the weights of 16-dimensional features were zero, which is equivalent to discarding the neurons in the locations of the 16-dimensional features, during the process of feature selection. In addition, for reducing the training time and preventing over-fitting, the "dropout" method is often used to temporarily discard certain neurons. However, in the "dropout" process where neurons are discarded randomly, some neurons with important characteristics may be discarded, leading to certain blindness, and in the process of training, it is necessary to eliminate neurons. It is difficult to debug the super-parameter of the quantity in training. In the process of feature selection, some neurons can be selectively discarded, based on the feature importance, playing a role not only in evaluating the importance of the features but also in selectively discarding neurons.

To verify the effectiveness of this feature selection method and the proposed defect recognition method, we consider CNN-1, the CNN architecture in Fig. 5 referred to as CNN-4, and CNN-5 as comparison experimental objects; more information on CNN-1, CNN-4, and CNN-5 is listed in Table 6. Under different iterations, experiments were carried out on CNN-1, CNN-4, and CNN-5, respectively, and the recognition accuracies are shown in Table 7. By comparing the accuracies of CNN-1 and CNN-4, the effectiveness of the proposed enhanced feature selection method can be verified. By comparing the accuracy of CNN-5 with those of CNN-1 to CNN-4, the validity of the welding defect recognition method proposed in this study can be verified.

The experiments demonstrate that the recognition accuracy can be further improved by combining the ReliefF algorithm with the neural network for feature selection. Moreover, when the number of iterations is relatively small, the advantage of this feature selection method is more obvious.

In Table 7, by comparing the defect recognition accuracy rates of CNN-1 and CNN-4 and CNN-3 and CNN-5 under different iterations, it can be seen that the method of enhanced feature selection proposed in this study can effectively improve the defect recognition rate, and when the number of iterations is 200, the defect image recognition accuracy rate of CNN-4 is 87.5%, which is 2.3% higher than that of CNN-1, and the defect image recognition accuracy of CNN-5 is accurate. The rate is 91.0%, which is 1.1% higher than that of CNN-3. The results show that the combination of ReliefF algorithm and neural network can further improve the recognition accuracy. Moreover, when the number of iterations is relatively small, the advantage of this enhanced feature selection method is more obvious.

By comparing the defect recognition accuracy of CNN-5 with CNN-1, CNN-3, and CNN-4 under different

Table 4 Corresponding weights of the initial feature value vector

Feature number							
1–7	0.0000	0.0758	0.0395	0.0140	0.0901	0.0358	0.0000
8–14	0.0000	0.0476	0.0000	0.0829	0.0395	0.0750	−0.0010
15–21	0.0333	−0.0023	−0.0001	0.0201	−0.008	0.0209	0.0132
22–28	0.1047	0.0104	0.0022	0.0431	0.1046	0.0686	0.0078
29–35	0.0263	0.0714	0.0455	0.0244	0.0400	0.1212	0.0000
36–42	0.0000	0.0288	0.0000	−0.0001	0.0065	0.0298	0.0528
43–49	0.0164	0.0004	0.0000	0.0339	0.0017	0.0605	−0.0007
50–56	0.0412	0.0329	0.0503	0.0426	0.0353	0.1215	−0.0044
57–63	0.1076	0.0724	0.0565	0.0019	0.0463	0.0365	0.0359
64–70	0.0337	0.2809	0.0711	0.0000	0.0037	0.0000	0.0756
71–77	0.0045	0.1191	0.0472	0.0914	0.0360	0.0013	0.0548
78–84	0.0212	0.0102	0.0350	0.0239	0.0153	0.1033	0.0693

Table 5 Corresponding weights of the revised feature value vector

Feature number							
1–7	0.0000	0.0758	0.0395	0.0140	0.0901	0.0358	0.0000
8–14	0.0000	0.0476	0.0000	0.0829	0.0395	0.075	0.0000
15–21	0.0333	0.0000	0.0000	0.0201	0.0000	0.0209	0.0132
22–28	0.1047	0.0104	0.0022	0.0431	0.1046	0.0686	0.0078
29–35	0.0263	0.0714	0.0455	0.0244	0.0400	0.1212	0.0000
36–42	0.0000	0.0288	0.0000	0.0000	0.0065	0.0298	0.0528
43–49	0.0164	0.0004	0.0000	0.0339	0.0017	0.0605	0.0000
50–56	0.0412	0.0329	0.0503	0.0426	0.0353	0.1215	0.0000
57–63	0.1076	0.0724	0.0565	0.0019	0.0463	0.0365	0.0359
64–70	0.0337	0.2809	0.0711	0.0000	0.0037	0.0000	0.0756
71–77	0.0045	0.1191	0.0472	0.0914	0.0360	0.0013	0.0548
78–84	0.0212	0.0102	0.0350	0.0239	0.0153	0.1033	0.0693

iterations, it can be seen that the model based on pooling strategy and feature selection has good recognition performance. When the number of iterations is 200, the recognition accuracy rate of defect image of the model is 91.0%, which is 1.1% higher than that of the model without enhanced feature selection, and compared with the model without enhanced feature selection. Compared with the traditional CNN model, the maximum pooling strategy is improved by 4.0% and 6.4%, respectively. Experimental results show that the proposed molten pool strategy and feature selection method have good effect on improving the defect recognition rate, and when the two are combined, the recognition accuracy of weld image defects can be further improved.

4.3 Validation of the proposed method

Sections 4.1 and 4.2 verified the proposed pooling method and feature selection method, respectively. In this section, the entire improved CNN is tested and validated. The data set used is the experimental data set of this study. The proposed method, CNN, SVM, and improved DS [10] method were used to perform the experiments, respectively. Table 8 shows the recognition accuracy for different defects under the various methods.

The proposed method can determine the category of defects in the input image. It can be seen from the table that the average accuracy of this method for defect identification is the highest among all the methods (7.57% higher than the DS method, 2.74% higher than the SVM method, and 3.29% higher than the CNN method).

Table 6 Model construction

CNN model	Pooling model		Add feature selection layer? (yes/no)
	P2 layer	P4 layer	
CNN-1	Max pooling	Max pooling	No
CNN-4	Max pooling	Max pooling	Yes
CNN-3	Proposed pooling strategy		No
CNN-5	Proposed pooling strategy		Yes

Table 7 Recognition rates of CNN-1 and CNN-4 under different iterations

Model	Times						
	Accuracy						
	50	75	100	125	150	175	200
CNN-1	0.828	0.846	0.846	0.864	0.855	0.864	0.855
CNN-4	0.848	0.866	0.864	0.884	0.875	0.884	0.875
CNN-3	0.866	0.866	0.875	0.884	0.893	0.900	0.900
CNN-5	0.875	0.875	0.884	0.893	0.893	0.900	0.910

Table 8 Comparison of the recognition accuracy of different methods (%)

Target class	Correct recognition rate of classification			
	Proposed method	CNN	SVM	DS
Accuracy	*91.71*	88.42	88.97	84.14
PO(32)	*97.56*	92.58	91.46	86.71
SL(41)	*96.78*	93.47	90.36	69.56
LP(38)	*92.94*	89.25	87.68	75.62
LF(35)	89.78	87.34	87.23	*90.56*
CR(26)	81.49	79.45	88.14	*98.26*

The maximum value of recognition accuracy of different methods for the same defect is marked in italics

With respect to the recognition accuracy of a single defect, the deep learning method applied by the proposed method achieves better results in identifying PO and SL defects due to its ability to extract abstract features; however, due to the small crack width, the number of features extracted by deep learning is less, and the recognition accuracy for CR defects needs to be improved. For traditional defects, the DS shows better performance. The results indicate that DS method has a high recognition accuracy for CR, but the accuracy is not high when identifying PO and SL defects. Further improvement involves the addition of artificial crack features to the CNN to improve the accuracy of crack identification.

5 Conclusion

In this study, in order to improve the pool adaptive ability and feature selection ability of the CNN for different defect image features, the classic pooling strategy was improved, and the traditional feature evaluation method was combined with the neural network for feature selection. In summary,

(1) A pooling strategy, which considers the feature distribution of the pooling region and the feature map to which the pooling region belongs, was proposed. This model includes the characteristics of max pooling and average pooling, and reflects the pooling region, when different feature distributions are involved. A certain degree of dynamic adaptability is significant for improving the recognition rate of deep neural networks.

(2) Combining the traditional feature evaluation method of ReliefF and the understanding of the feature importance of the neural network, the feature selection ability of the model was strengthened, enabling further improvement of the model's classification ability.

(3) The method proposed in this study can identify and classify defects in radiographic images. The effectiveness of the CNN model based on the improved pooling strategy

and feature selection was verified. The experimental results demonstrated that compared to the traditional CNN, the proposed method has higher correct recognition rate and better adaptability. Compared to the traditional DS method, the overall performance of the proposed method was improved; however, the recognition accuracy for crack defects requires improvement. The CNN model based on the improved pooling strategy and feature selection exhibited good performance in the defect classification of X-ray images. In the future, it is intended to improve the recognition accuracy for crack defects.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Zhang Z, Chen S (2017) Real-time seam penetration identification in arc welding based on fusion of sound, voltage and spectrum signals[J]. J Intell Manuf 28(1):207–218. https://doi.org/10.1007/s10845-014-0971-y
2. He K, Li X (2016) A quantitative estimation technique for welding quality using local mean decomposition and support vector machine[J]. J Intell Manuf 27(3):525–533. https://doi.org/10.1007/s10845-014-0885-8
3. Chady T, Sikora R, Misztal L, Grochowalska B, Grzywacz B, Szydłowski M, Waszczuk P, Szwagiel M (2017) The application of rough sets theory to design of weld defect classifiers. J Nondestruct Eval 36(2):40. https://doi.org/10.1007/s10921-017-0420-x
4. Liao Z, Sun J (2013) Image segmentation in weld defect detection based on modified background subtraction. Proceedings of the 2013 6th International Congress on Image and Signal Processing (CISP): 610-615. https://doi.org/10.1109/CISP.2013.6745239
5. Lim TY, Ratnam MM, Khalid MA (2007) Automatic classification of weld defects using simulated data and an MLP neural network. Insight Non-Destruct Test Cond Monit 49(6):154–159. https://doi.org/10.1784/insi.2007.49.3.154
6. Nacereddine N, Ziou D, Hamami L (2013) Fusion-based shape descriptor for weld defect radiographic image retrieval. Int J Adv Manuf Technol 68(9–12):2815–2832. https://doi.org/10.1007/s00170-013-4857-5
7. Jiang H, Zhao Y, Gao J, Wang Z (2016) Weld defect classification based on texture features and principal component analysis. Insight Non-Destruct Test Cond Monit 58(4):194–200. https://doi.org/10.1784/insi.2016.58.4.194
8. Goumeidane AB, Bouzaieni A, Nacereddine N et al (2015) In: Azzopardi G, Petkov N (eds) Bayesian networks-based defects classes discrimination in weld radiographic images. Computer Analysis of Images and Patterns- Proceedings of 16th

International Conference (CAIP), vol 9257, Valletta, pp 554–565. https://doi.org/10.1007/978-3-319-23117-4_48

9. Mekhalfa F, Nacereddine N (2014) Multiclass classification of weld defects in radiographic images based on support vector machines. Proceedings of Tenth International Conference on Signal-Image Technology and Internet-Based Systems (SITIS):1-6. https://doi.org/10.1109/SITIS.2014.72

10. Jiang H, Wang R, Gao Z, Gao J, Wang H (2019) Classification of weld defects based on the analytical hierarchy process and Dempster–Shafer evidence theory. J Intell Manuf 30(4):2013–2024. https://doi.org/10.1007/s10845-017-1369-4

11. Lu H, Zhang M, Liu Y, Ma S (2017) Convolution neural network feature importance analysis and feature selection enhanced model. J Softw 28(11):2879–2890 (in Chinese). https://doi.org/10.13328/j.cnki.jos.005349

12. Juefei-Xu F, Boddeti V N, Savvides M (2017) Local binary convolutional neural networks. Proceedings of 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017):4284-4293. https://doi.org/10.1109/CVPR.2017.456

13. Tokozume Y, Harada T (2017) Learning environmental sounds with end-to-end convolutional neural network. Proceedings of 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP):2721-2725. https://doi.org/10.1109/ICASSP.2017.7952651

14. Ghazaei G, Alameer A, Degenaar P et al (2017) Deep learning-based artificial vision for grasp classification in myoelectric hands. J Neural Eng 14(3):036025. https://doi.org/10.1088/1741-2552/aa6802

15. Liu K, Niu Y, Yang J et al (2016) Product related information sentiment-content analysis based on convolutional neural networks for the Chinese micro-blog. Proceedings of 2016 International Conference on Network and Information Systems for Computers (ICNISC):357-361. https://doi.org/10.1109/ICNISC.2016.083

16. He K, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Trans Pattern Anal Mach Intell 37(9):1904–1916. https://doi.org/10.1109/TPAMI.2015.2389824

17. Zhang Z, Chen D, Wang J, Bai L, Hancock ER (2018) Quantum-based subgraph convolutional neural networks. Pattern Recogn 88:38–49. https://doi.org/10.1016/j.patcog.2018.11.002

18. Suganuma M, Shirakawa S, Nagao T (2017) A genetic programming approach to designing convolutional neural network architectures. Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'17):497-504. https://doi.org/10.1145/3071178.3071229

19. Yan Z, Zhang H, Piramuthu R, et al (2016) HD-CNN: hierarchical deep convolutional neural networks for large scale visual recognition. Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV):2740-2748. https://doi.org/10.1109/ICCV.2015.314

20. Kokkinos I (2017) UberNet: training a universal convolutional neural network for low-, mid-, and high-level vision using diverse datasets and limited memory. Proceedings of 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR):5454-5463. https://doi.org/10.1109/CVPR.2017.579

21. Chéron G, Laptev I, Schmid C (2015) P-CNN: pose-based CNN features for action recognition. Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV): 3218-3226. https://doi.org/10.1109/ICCV.2015.368

22. Chaturvedi I, Cambria E, Poria S et al (2016) Bayesian deep convolutional belief networks for subjectivity detection. Proceedings of 16th IEEE International Conference on Data Mining Workshops (ICDMW):916-923. https://doi.org/10.1109/ICDMW.2016.0134

23. Ha M, Byeon Y, Lee Y et al (2019) Selective deep convolutional neural network for low cost distorted image classification. IEEE Access 7:133030–133042. https://doi.org/10.1109/ACCESS.2019.2939781

24. Wang J, Yang Y, Mao J et al (2016) CNN-RNN: a unified framework for multi-label image classification. Proceedings of 29th IEEE Conference on Computer Vision and Pattern Recognition (CVPR):2285-2294. https://doi.org/10.1109/CVPR.2016.251

25. Wu X, He R, Sun Z, Tan T (2018) A light CNN for deep face representation with noisy labels. IEEE Trans Inf Forensics Secur 13(11):2884–2896. https://doi.org/10.1109/TIFS.2018.2833032

26. Alom MZ, Hasan M, Yakopcic C (2020) Improved inception-residual convolutional neural network for object recognition. Neural Comput & Applic 32(1):279–293. https://doi.org/10.1007/s00521-018-3627-6

27. Abdulnabi AH, Wang G, Lu J, Jia K (2015) Multi-task CNN models for attribute prediction. IEEE Trans Multimedia 17(11):1949–1959. https://doi.org/10.1109/TMM.2015.2477680

28. Chen J, Patel VM, Chellappa R (2016) Unconstrained face verification using deep CNN features. 2016 IEEE Winter Conference on Applications of Computer Vision (WACV):7477557. https://doi.org/10.1109/WACV.2016.7477557

29. Khumaidi A, Yuniarno EM, Purnomo MH (2017) Welding defect classification based on convolution neural network (CNN) and Gaussian kernel. Proceedings of 2017 International Seminar on Intelligent Technology and Its Applications (ISITIA):261-265. https://doi.org/10.1109/ISITIA.2017.8124091

30. Liu B, Zhang X, Gao Z et al (2017) In: Zhai G, Zhou J, Yang X (eds) Weld defect images classification with VGG16-based neural network. Communications in Computer and Information Science - Proceedings of 14th International Forum on Digital TV and Wireless Multimedia Communications (IFTC), Singapore, pp 215–223. https://doi.org/10.1007/978-981-10-8108-8_20

31. Kumar et al (2020) Improved deep convolution neural network (CNN) for the identification of defects in the centrifugal pump using acoustic images. Appl Acoust 167. https://doi.org/10.1016/j.apacoust.2020.107399

32. Yuan et al (2020) Research on pipe surface defect recognition based on convolutional neural network. Modern Electron Tech:47–51. https://doi.org/10.16526/j.cnki.11-4762/tp.2020.06.038

33. Xie et al (2020) Classification and recognition method of sheet metal parts surface defects based on convolution neural network. Comput Meas Control:187–190+196. https://doi.org/10.16526/j.cnki.11-4762/tp.2020.06.038

34. Fan D et al (2020). X-ray image defect recognition method for pipe weld based on improved convolutional neural network. Hanjie Xuebao,7-11. https://doi.org/10.12073/j.hjxb.20190703002

35. Miao et al (2019) Online defect recognition of narrow overlap weld based on two-stage recognition model combining continuous wavelet transform and convolutional neural network. Comput Ind:112. https://doi.org/10.1016/j.compind.2019.07.005

36. Li et al (2019) Recognition of X-ray weld defects based on deep learning. J Xi'an Shiyou Univ (Nat Sci Ed):74–81. https://doi.org/10.3969/j.issn.1673-064X.2019.04.012

37. Tareef A, Song Y, Huang H, Wang Y, Feng D, Chen M, Cai W (2017) Optimizing the cervix cytological examination based on deep learning and dynamic shape modeling. Neurocomputing 248:28–40. https://doi.org/10.1016/j.neucom.2017.01.093

38. Gore S, Govindaraju V (2016) In: Skulimowski A, Kacprzyk J (eds) Feature selection using cooperative game theory and relief algorithm. advances in intelligent systems and computing - knowledge, information and creativity support systems: recent trends, advances and solutions (KICSS), Cham, pp 401–412. https://doi.org/10.1007/978-3-319-19090-7_30

39. Reyes O, Morell C, Ventura S (2015) Scalable extensions of the ReliefF algorithm for weighting and selecting features on the multi-label learning context. Neurocomputing 161:168–182. https://doi.org/10.1016/j.neucom.2015.02.045