




Recognition of weld defects from X-ray images based on improved convolutional neural network

Ande Hu¹ · Lijian Wu² · Jiankang Huang²  · Ding Fan¹ · Zhenya Xu³

Received: 18 July 2020 / Revised: 24 January 2022 / Accepted: 31 January 2022 /
Published online: 26 February 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

When convolutional neural network (CNN) is used for welding defect detection image recognition, the recognition result will be affected by many factors such as human factors, the activation function is sensitive to input parameters, and the edge features are weakened. In order to overcome the above problems, the methods include image processing, exponential linear unit (ELU) activation function and improved pooling model are used. According to the experiment, the image processing method can effectively segment the weld and defects, and the defect location in the weld image can be located. Using the ELU activation function in the CNN model can improve the robustness of the neural network to the input parameters and increase the sparsity of the network to increase the model's convergence speed. The improved pooling method based on grayscale adaptation can increase the extraction range of weld defect features and reduce the impact of noise, and has certain dynamic adaptability to the defect features. The result shows that the improved convolutional neural network(ICNN) method can effectively improve the accuracy of recognition in weld image recognition, and the overall recognition rate can reach 98.13%.

Keywords Weld defect recognition · Convolution neural network · ELU function · Pooling

✉ Jiankang Huang
sr2810@163.com

✉ Ding Fan
fand@lut.cn

¹ State Key Laboratory of Advanced Processing and Recycling of Non-ferrous Metals, Lanzhou University of Technology, Lanzhou 730050, China

² School of Materials Science and Engineering, Lanzhou University of Technology, Lanzhou 730050, China

³ Baoshan Iron and Steel Co., Ltd, Shanghai 201900, China

1 Introduction

In the processing of industrial welding production, various defects will be generated due to the instability of weld parameters, which are mainly divided into external defects (undercut, excess weld metal, cavities, etc.) and internal defects (crack, gas pore, slag inclusions, lack of penetration, etc.) [20]. X-ray inspection technology is often used to detect internal defects in welds, and the results are identified and evaluated manually. Given the problems such as misdetection, false detection, and inefficiency in the process of current manual inspection and evaluation, it is particularly important to seek an automatic detection and recognition method for welding defects, which can make defect detection efficient, standardized and intelligent [19].

CNN does not need to manually describe and extract the target image, it can learn the features from the training samples autonomously through the neural network, and these features are closely related to the classifier, which solves the problem of manually extracting features and classifier selection. Meanwhile, using its 'end-to-end' advantage to solve some problems that were considered difficult to solve in the past defect detection of weld images [23]. The widely used CNN model is the simplified version of the Hubel-Wiesel model [30], and the research on this model mainly focused on algorithm improvement [8, 11, 13, 27] and structural improvement for different fields [6, 12, 21, 22, 24]. The deep learning model improved by the algorithm has achieved significant results in image recognition, but there are still some shortcomings in the defect recognition of weld images based on x-rays, for example: the needed volume of image data during the training process is large, a small amount of data is easy to cause the neural network to underfit, while the number of industrial-level weld defect images can be used is small, which is difficult to meet the training requirements. At the same time, the human workload is large in image processing, and the defects need to be located manually. In addition, Rectified Linear Unit (ReLU) as a non-saturated activation function, there is a phenomenon of neuron death during training. When a large gradient flow through the neuron and the parameters are updated, the neuron will no longer activate. If there is a large learning rate, it will lead to excessive neuron death, which will affect the accuracy of training. The common pooling methods used in traditional convolutional neural network models are mean pooling and max pooling. In the weld images, for the grayscale changes at the weld zone, defect and the transition area between them, the defect characteristics may weaken if using mean pooling, the use of max pooling may introduce noise.

According to the shortcomings and deficiencies of the existing CNN model in weld flaw detection images, this paper uses an adaptive pooling layer and an improved activation function to test its improvement in weld flaw detection image recognition. Many existing studies have started with the recognition of the types of defects in the weld inspection image and the accuracy of the recognition. In terms of the recognition of the types of defects, the CNN model used in the article can identify five types of defects including no defects. In terms of comparison with other existing methods, it shows the superiority and robustness of this method in recognition. First, the image enhancement method is used to expand the existing industrial welding images to prevent insufficient fitting due to lack of training data. At the same time, the ELU improved activation function feature model selection method [1, 7] is used to identify weld defect detection image defects, and the effectiveness and correct rate of the improved activation function in weld defect detection image defect recognition are studied. The merging method will affect the weld image recognition. We use an improved merging method based on gray-scale adaptation, which can comprehensively consider the impact of

gray-level changes on the weld area, defects and transition areas. Compared with traditional methods, the gray-scale feature extraction from the input features has a certain dynamic adaptability.

2 X-ray image weld detection and defect extraction

Understanding the types of defects of tube welds is the basis for correct detection of tube weld defects. The formation mechanism of various weld defects is different, and the image features on X-ray images are also different in shape, but the main defects and their characteristics have the following types: (1) cracks(CK): cracks appear as white lines with irregular shapes and thickness on X-ray pictures, and can be divided into horizontal or vertical cracks according to shape; (2) gas pore(GP): according to the different forms of expression, it can be divided into single gas pore, chain gas pore and dense gas pore, etc. The image is white round or oval; (3) lack of penetration(LOP): the characteristic of X-ray image is a thin line of white piles with regular shape but irregular length, and the direction is generally along the direction of the weld bead; (4) lack of fusion(LOF): its X-ray image features appear as continuous or intermittent black lines; (5) flawless(FL): there are no obvious defects on the X-ray image. Figure 1 shows the X-ray images of each defect.

In X-ray weld images, the target area is relatively small compared to the overall image, redundant image information makes subsequent data processing and training more difficult, and various types of noise in the base material area will also have a greater impact. The image processing method can effectively segment the weld seam and the defect area, reduce the

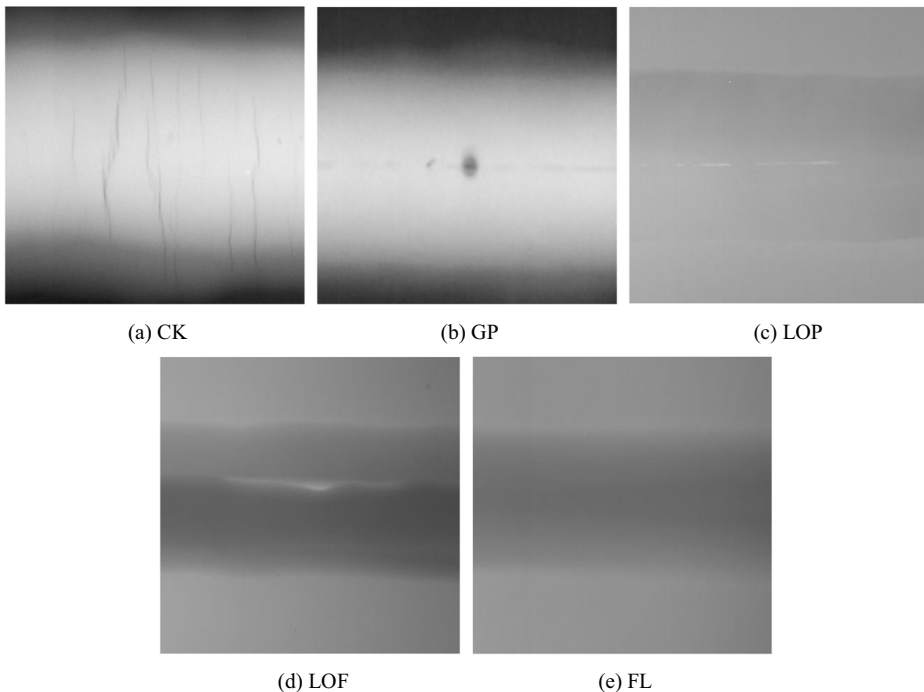


Fig. 1 X-ray weld defect images

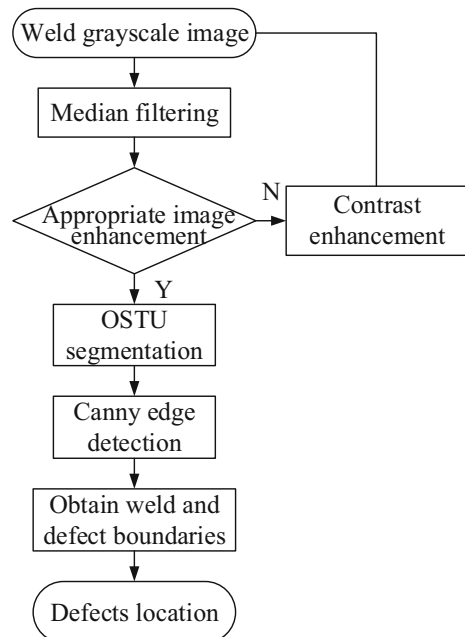
redundant information in the image, and prevent the influence of the base material on the result. Due to the different formation mechanisms of various types of weld defects, the imaging features on X-ray inspection images also have different shapes. However, the boundaries of the welds are approximately straight lines, and the difference in thickness between the weld and the base metal causes its gray value to change, so the weld and defects can be extracted. The X-ray image processing process is divided into two parts: welding seam detection and defect location. The processing flow is shown in Fig. 2.

2.1 Weld detection

In the process of X-ray inspection, because the thickness of the base metal, the weld, and the defects are all different, different gray areas appear in the digital image through exposure. Using this feature can achieve the purpose of detecting the weld area and extracting defects. The generation of X-ray image noise is related to the process of image formation and transmission channels, mainly including quantum noise caused by ray exposure, shot noise caused by irregular emission of electrons, etc. [2]. To reduce the influence of noise on welds and defects, filtering technology is used to remove noise. For the same weld flaw detection image, median filter, mean filter, Gaussian filter, and bilateral filtering are used to reduce noise. Figure 3 shows the peak signal to noise ratio (PSNR) scatter plot after reducing the noise of gas pore, cracks, lack of fusion, lack of penetration and defect less sample image. Through comparison, it can be seen that the median filtering technology can reduce the proportion of noise in the signal, so median filtering can be used to denoise X-ray images.

The denoised image is enhanced with a limiting linear stretching enhancement method, that is, the gray levels of the low gray value and high gray value pixel units of the image are appropriately combined, and the gray value of the middle part is stretched. The limiting linear

Fig. 2 The process of weld region extraction



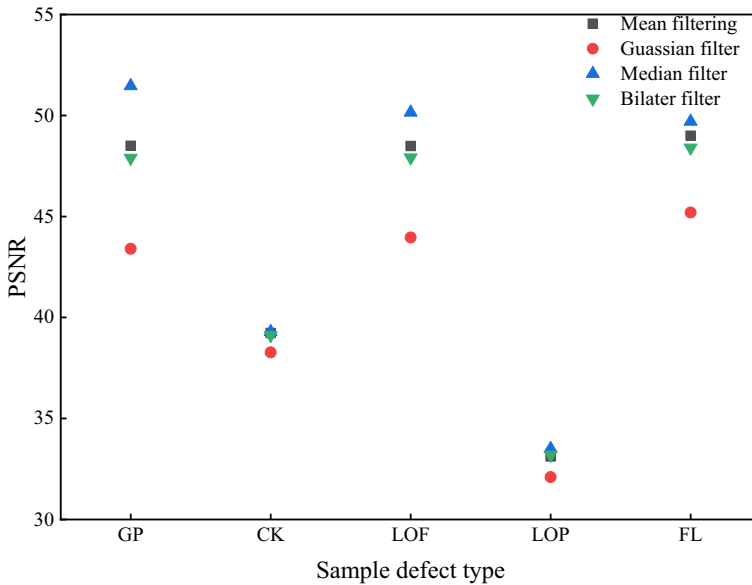


Fig. 3 PSNR diagrams of four different filtering methods

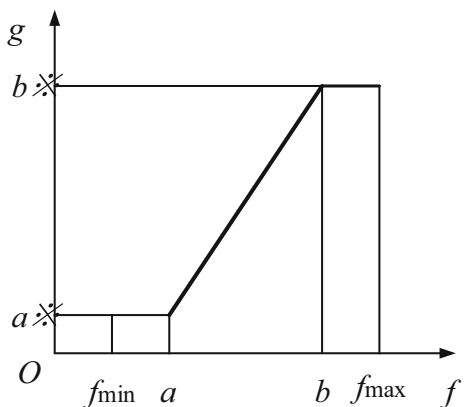
stretching of the image is shown in Fig. 4, the range of the limiting in the figure can be freely selected. The calculation formula is shown in Eq. (1).

$$g(x,y) = a' + \frac{b'-a'}{b-a} \times [f(x,y)-a] \tag{1}$$

Where $f(x, y)$ is the gray value at (x, y) , $g(x, y)$ is the gray value output through the mapping transformation, and a, b is the gray stretch interval. The sample image of the weld before and after the limiting linear stretching is shown in Fig. 5. It can be seen from the figure that the contrast of the image after the limiting linear stretching is more obvious, and the weld boundary and defects can be clearly distinguished.

There are differences in the gray distribution of the images after limiting linear stretching enhancement. Conventional image segmentation methods cannot well adapt to weld inspection

Fig. 4 Image linear stretching



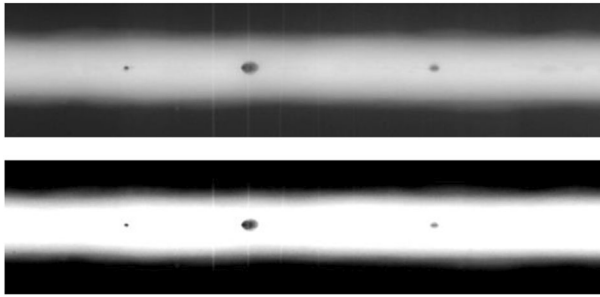


Fig. 5 Contrast enhancement

images with different gray distributions, the OSTU algorithm can be used to segment weld inspection images. The grayscale image is divided into two parts by the adaptive threshold K , namely the target A and the background B , so that the inter-class variance of A and B reaches the maximum. The definition of the inter-class variance between target A and background B is Eq. (2):

$$E^2(K) = P_a(\sigma - \sigma_a)^2 + P_b(\sigma - \sigma_b)^2 \quad (2)$$

Where σ is the gray value of the image, and σ_a and σ_b are the averages of target A and background B respectively, when $E^2(K)$ reaches the maximum value, K is the optimal threshold value. The binary image after the OSTU method is used to segment the weld region is shown in Fig. 6.

2.2 Weld defect location

The binary image can be used to directly distinguish the base metal, weld zone and defect position in the flaw detection negative film. To automatically locate the defect location in the weld zone, an edge detection method based on the canny operator [5] can be used. Different thresholds to detect strong edges and weak edges respectively, and only when strong edges and weak edges are connected, weak edges can be included in the image to avoid being filled with weak noise [14, 17]. During the canny edge detection process, the weld area and defects can be detected at the same time. The obtained weld and defect boundary information are shown in Fig. 7a. For different types of defects in the weld, the position coordinates of the weld defect can be obtained at the same time by using the 8-chain code boundary traversal tracking method, and the center position coordinates can be calculated by the horizontal and vertical

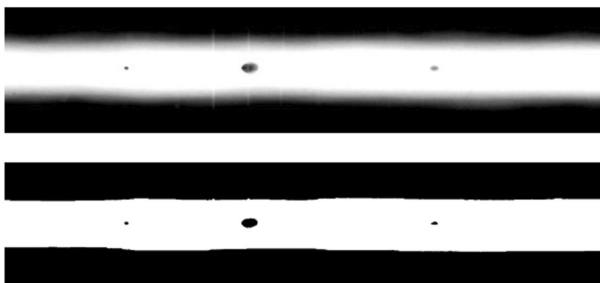


Fig. 6 OSTU binarized segmentation image

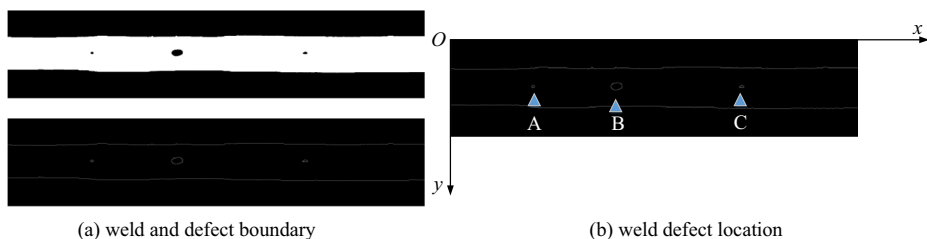


Fig. 7 Weld region extraction and defects location. **a** weld and defect boundary, **b** weld defect location

coordinate values, and finally to complete the weld defect Positioning. Using coordinates to locate the weld defect is shown in Fig. 7b. The gas pore defects contained in it include A, B, and C. The position coordinates of each point are: A (448, 255), B (899, 253), and C (1580, 247).

This section first introduces the image characteristics of the tube weld X-ray image, and then uses a series of image processing methods to process the weld flaw detection images, and extracts and locates the circular defect area, which verifies the reliability of the method.

3 Improved CNN feature selection method

This part first introduces the basic structure of the traditional convolutional neural network model. Secondly, through the comparison between the traditional activation function and the improved activation function principle, it shows the superiority of the improved activation function in the recognition of weld defects. Finally, the principle of adaptive pooling method and its role in image recognition are introduced.

3.1 CNN network overview

The basic structure of a CNN consists of an input layer, a convolutional layer, a pooling layer, a fully connected layer, and an output layer. In the CNN model, the convolutional layer and the pooling layer include multiple layers and are alternately connected. A typical CNN model is shown in Fig. 8.

It can be seen that each neuron of the output feature surface in the convolutional layer is locally connected to its input [9], and each weighted sum of the linear function of the weight and bias value is passed to a non-linear function such as the ReLU function to obtain the output value of each neuron, this process is the convolution process [10]. Hidden layers containing multiple convolutional layers and down-sampling layers in the convolutional neural network can perform feature extraction to obtain feature vectors of a certain dimension, and obtain the output of classification results through the output layer. In addition, a weight sharing strategy can also be used and it can effectively reduce the number of parameters to be trained in the neural network and increase the training speed. After multi-layer convolution pooling, one or more fully connected layers are usually connected, and each neuron in the fully connected layer is fully connected to all neurons in the previous layer [28, 29]. In order to improve the performance of the CNN network, the excitation function of each neuron in the fully connected layer generally uses the ReLU function [16], and the output layer is classified by softmax regression.

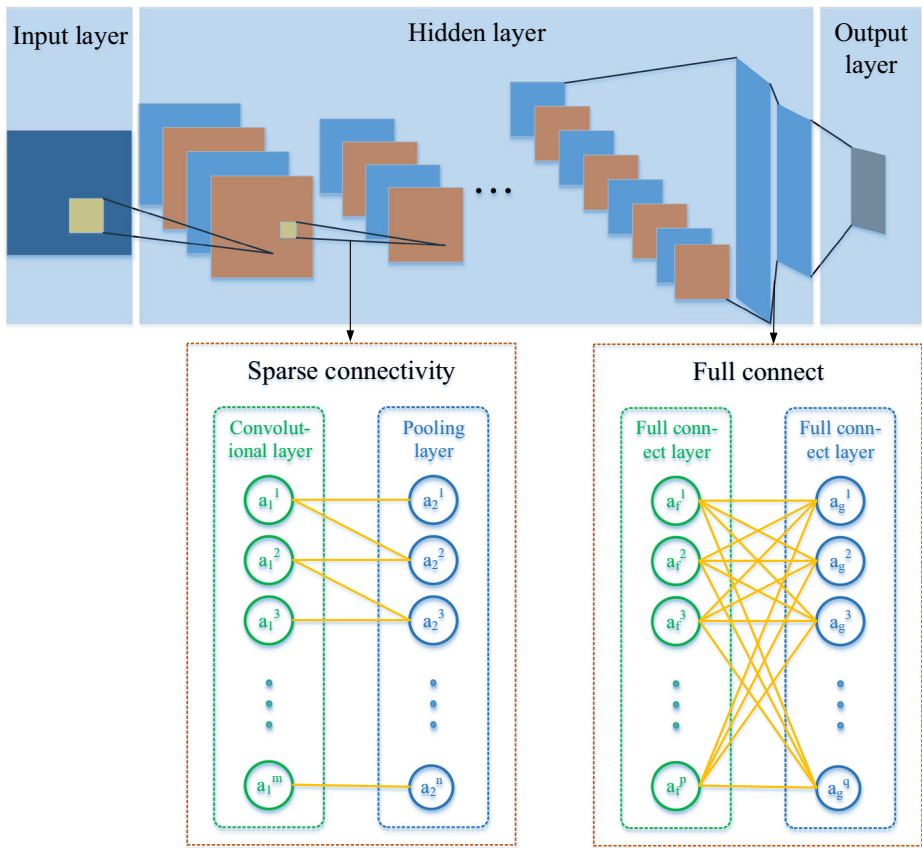


Fig. 8 Typical convolutional neural network

3.2 ELU activation function application

For the neural network, the activation function can be used to introduce nonlinear transformation to the neuron, and the neural network can be approximated by any nonlinear function through training, then the different features of the function in the network recognition can be fitted. The sigmoid function has an exponential function shape, which is similar to biological neurons and is widely used in artificial neural networks. However, its output in the x direction gradually approaches zero and has soft saturation, once it enters the saturation region during training, appearing gradient disappearance will make it difficult to effectively train the network parameters. The unilateral suppression ability of the ReLU activation function can make the neurons in the network sparsely activating, and thus better mine relevant features and fit the training data, effectively solving the gradient explosion/gradient disappearance problem. It can be seen that when $x > 0$, there is no saturation problem in the ReLU function, and the gradient can be kept attenuated during the training process, thereby alleviating the problem of gradient disappearance. When $x < 0$, hard saturation occurs. As the training progresses, some of the inputs enter the hard saturation zone, causing the weights to be unchanged. Therefore, the ReLU activation function is weak in the network training process, and the problem that the

neurons no longer react when the large gradient flows through the neurons and updates the parameters.

For the problems in traditional convolutional neural networks, this paper adopts an ELU nonlinear activation function that can comprehensively consider the saturation of the activation function, the expression is shown in (3).

$$f(x) = \begin{cases} x & x \geq 0 \\ a(e^x - 1) & x < 0 \end{cases} \tag{3}$$

From this expression we can see that the activation function can fuse the advantages of the Sigmoid and ReLU functions and maintain the unsaturation on the right side of the function. At the same time, the soft saturation on the left side of the function is increased, so that the non-saturation part can alleviate the gradient disappearance phenomenon during the model training process, the soft saturation can make the model more robust to the input parameter and existing noise [3, 4, 25]. Various activation function images are shown respectively in Fig. 9.

3.3 Pooling method based on gray feature adaptive

For the problems in the weld inspection image of the pooling method selected in the traditional convolutional neural network, an adaptive pooling method that comprehensively considers the defects and the gray level of the weld zone is used to characterize the pooling domain and feature maps. The values are dynamically adjusted to prevent feature weakening and noise effects. The principle of the improved pooling method based on adaptive grayscale features is shown in Fig. 10.

During the pooling process, the area where the local defect is located abstracted as a pooling domain, the feature layer of the convolution layer is abstracted as a feature map, the relationship between the feature variance σ_p in the pooling domain and the feature map variance σ_{FI} is used to construct a correction factor μ for the pooling domain. Eigenvalues with large variances are modified to improve the traditional average pooling model, the expression is shown in Eq. (4).

$$S = \begin{cases} \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n F_{ij}, \sigma_p < \sigma_{FI} \\ \frac{\mu}{n^2} \left(\sum_{i=1}^n \sum_{j=1}^n F_{ij} \right), \sigma_p \geq \sigma_{FI} \end{cases} \tag{4}$$

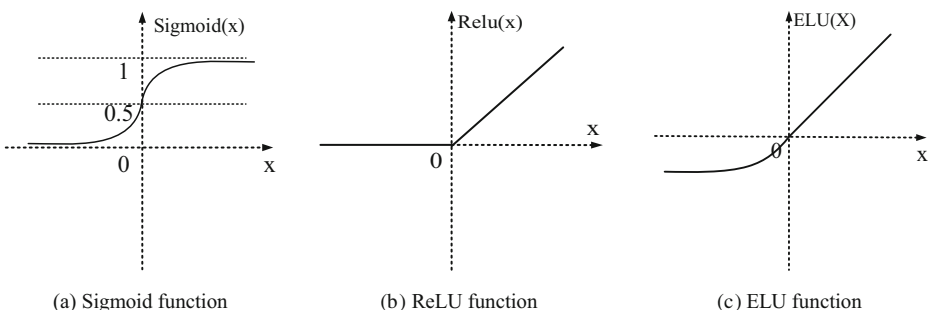


Fig. 9 activation functions. a Sigmoid function, b ReLU function, c ELU function

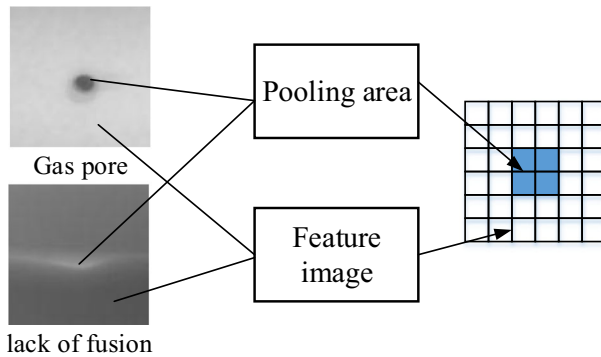


Fig. 10 Improved pooling principle abstraction

When $\sigma_p < \sigma_{FI}$, the feature distribution in the pooling domain is uniform, and the feature extraction method is the same as the average pooling method. When $\sigma_p \geq \sigma_{FI}$, the feature variance of the pooling domain is large. At this time, a correction factor μ is introduced to modify the average pooling model, the calculation method of the correction factor μ is shown in (5):

$$\mu = \frac{p_{sum}}{p_{sum} - (p_{max} - p_{min})} \quad (5)$$

Among them, p_{sum} , p_{max} , p_{min} are the eigenvalue sum, the maximum and minimum eigenvalues in the pooling domain respectively. The pooling model can comprehensively consider the gray value changes of the welding defect and the transition area around the defect, improve the feature extraction area of the weld defect, and have a certain dynamic adaptability for feature extraction at different positions.

4 Improved CNN model construction and defect recognition

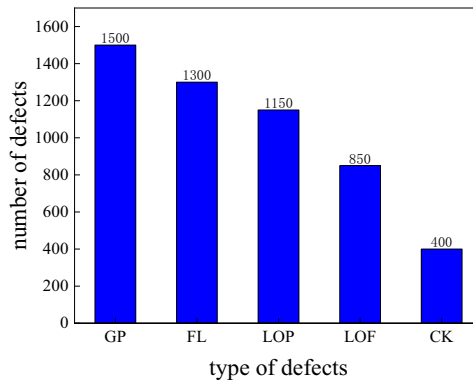
This part mainly introduces the collection and processing methods of training and verification data sets used to improve the model, the training and processing procedures of different models, and the comparison of the final conclusions.

4.1 Image data processing of weld defects

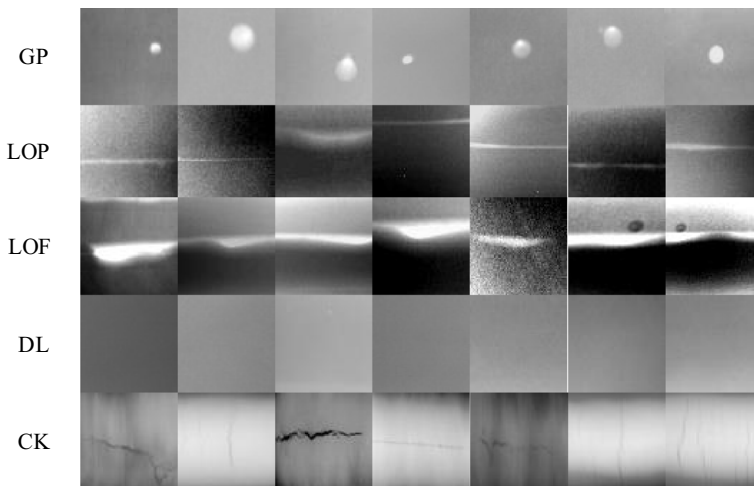
The radiographic image of the weld selected in this paper consists of two parts, one is taken from a public database named GDXray [15], and the other is a radiographic image of a pipe weld provided by a domestic welding processing company. Because the area of interest of the image required during image processing is small, and the proportion of uninteresting areas in the weld inspection image as a whole is relatively large, it is difficult to use the overall image for training. Therefore, before the CNN model is trained, image pre-processing is performed. According to the type and size of the defect of the weld flaw detection image, the OpenCV image processing tool is used to normalize the five types of weld flaw detection images to 68×68 size area of interest. In order to further increase the amount of training data for weld inspection images, image enhancement technology was used to expand the original data image

by 1:10, and the final sample was expanded to 5200, including gas pore(GP), crack(CK), defectless(DL), lack of penetration(LOP), lack of fusion(LOF). Manual classification method is used to store the images in 5 folders according to the type of defect, and the label is 1–5, and the naming format of the image is image_x(y), where x is the defect label and y is the type of defect Sample serial number, and finally divide all data into training set, verification set and test set according to 8:1:1. The number of samples of various defects and some experimental image data is shown in Fig. 11a, b.

In the image processing process, the input image samples are first shuffled and arranged randomly to form an input file queue, then the file is subjected to operations such as image widening and enhancement, and finally a data combination queue is formed, which is input into the deep learning framework input layer. Input image processing flow is shown in Fig. 12.



(a) experimental image data statistics



(b) Partial experimental image data

Fig. 11 Number of various types of defects and partial experimental image data. **a** experimental image data statistics, **b** Partial experimental image data

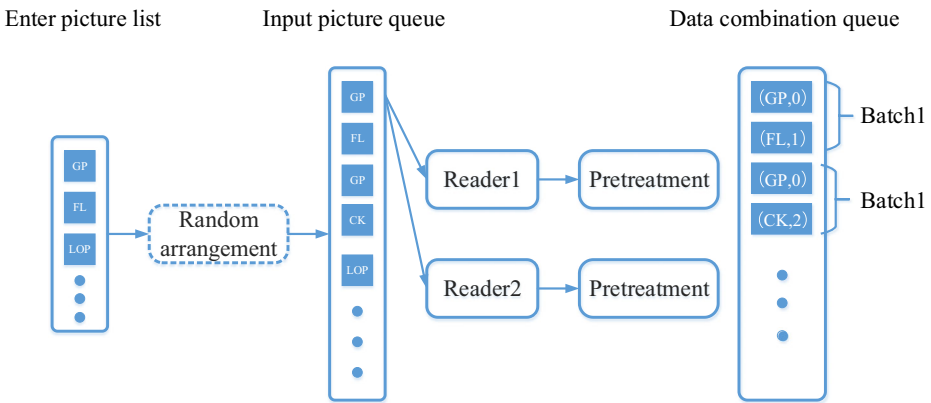


Fig. 12 Input image processing flow

4.2 Improved CNN model construction

For the convolutional neural network model, the connection layers and depth of the convolutional neural network can be adjusted appropriately according to the size of the input layer. As the model depth increases, the learning effect is also better, but increasing the depth of the network will increase the calculation time and network parameter, and if the training data is insufficient will increase the risk of overfitting, so in the experiment increasing the depth of the network layer is not the first choice for the network model. In the establishment of the convolutional neural network model, the network parameters are effectively selected to obtain the maximum output with the minimum number of layers. By connecting the local receptive fields of each feature surface, the original pixels of the input image are mapped hierarchically to extract the layers of the receptive fields. Using weight sharing strategies to reduce the amount of data in the neural network, and additionally change the activation function to reduce the complexity of the model and make the network easier to train. To verify the effectiveness of the method used in this paper, models named CNN-1, CNN-2, CNN-3, CNN-4 and CNN-5 were constructed. The CNN-1 model uses the ReLU activation function, and the CNN-2 model uses the ELU activation function, all pooling layers use the mean pooling method. CNN-3 model uses maximum pooling, and CNN-4 uses the improved pooling method described in this article, the activation function uses the ReLU function. The construction methods of each model are shown in Table 1. By constructing different CNN models for comparative experiments, CNN-1 and CNN-2 model tests use the same pooling method to verify the validity of the ELU activation function; CNN-1, CNN-2, and CNN-3 model tests all use ReLU activation Function to verify the effectiveness of the improved pooling method in

Table 1 Structure method of CNN model

Model name	Activation function	Pooling method
CNN-1	ReLU	Mean-pooling
CNN-2	ELU	Mean-pooling
CNN-3	ReLU	Max-pooling
CNN-4	ReLU	Improved-pooling
CNN-5	ELU	Improved-pooling

this article by using different pooling methods; CNN-5 model can compare the recognition rate under the condition of ELU activation function and improved pooling model.. The improved CNN model is shown in Fig. 13.

The input image is the input layer and C is the convolution layer. The size of the convolution kernel is 5×5 , and the depth is 6, 12, and 16 in turn. Each convolution layer consists of several convolution units and uses the back-propagation algorithm to optimize the parameters of the convolution unit, iteratively extract more complex features by extracting different features of the input. N is regularization, which can constrain the convolution results. E is the ELU activation function, the calculation result is de-linearized by the ELU function. P is the pooling layer, the size of the convolution kernel is set to 2×2 , the moving step is 2. The convolution layer and the pooling layer are all filled with 0. FC is a fully connected layer. Through two fully connected layers, the number of nodes can reduce to 60 through. Because the type of CNN to be classified and identified in this paper is 5 types, the number of output layers S is set to 5.

4.3 Image defect recognition and analysis

The experiment is based on the Linux Ubuntu16.04 operation system and is performed under the Tensorflow framework. CNN-1, CNN-2, CNN-3, and CNN-4 models are trained 400 times using the image data set provided in the article. Using the Tensorboard visualization module to visualize the accuracy and cross-entropy loss during the iteration process and store the data at each step. After the iteration is completed, the data can visualized and output. The training accuracy rate, verification accuracy rate and cross-entropy loss, which was shown in Fig. 14. By comparing and analyzing the CNN-1 and CNN-2 training accuracy and cross-entropy loss changes in Fig. 14a, b, we can see that under the same number of iterations, the CNN model with the ELU activation function convergence faster than the ReLU activation function, and the training accuracy and cross-entropy loss value have stabilized at the 100th epoch, while the CNN model with the ReLU activation function has a slower convergence rate and reaches dynamic stability at 200th epoch. As is shown in Fig. 14c, It can be seen from the change of the correctness of the validation set that using the ELU activation function also has a better effect in terms of recognition accuracy. From this comparison test, it can be concluded that the ELU activation function has better convergence and better defect recognition ability

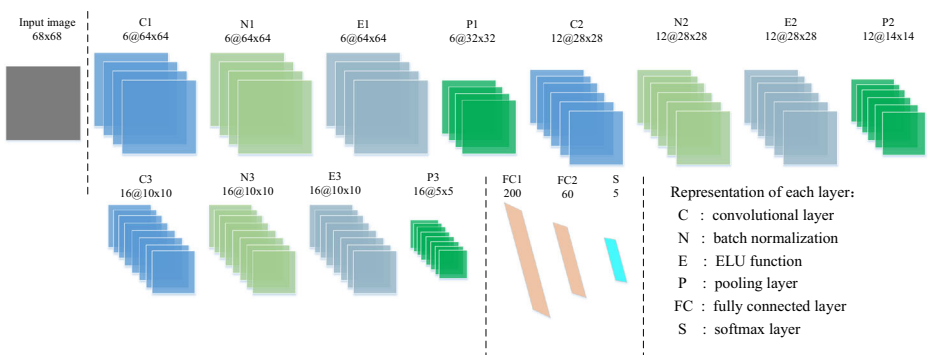
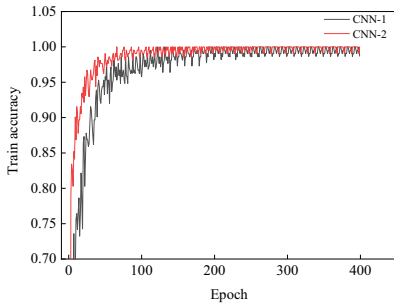
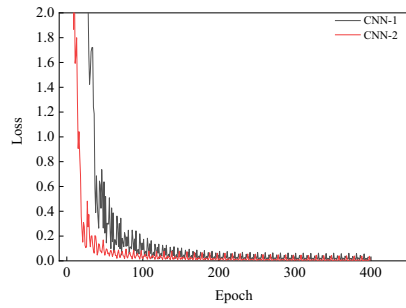


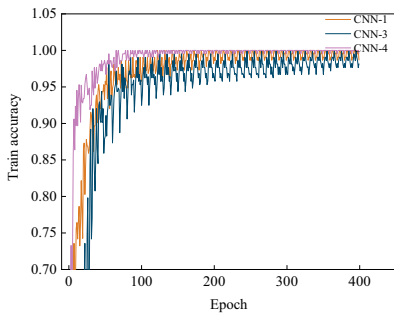
Fig. 13 Improved CNN model



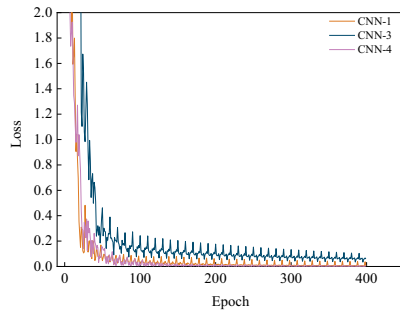
(a) the training accuracy rate of CNN-1 and CNN-2



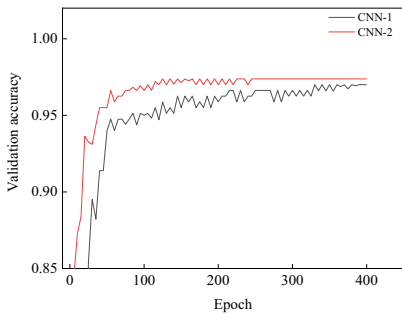
(b) the cross-entropy loss of CNN-1 and CNN-2



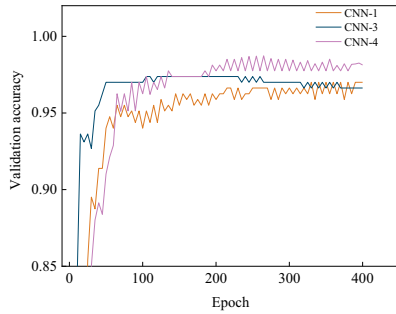
(c) the training accuracy rate of CNN-1, CNN-3 and CNN-4



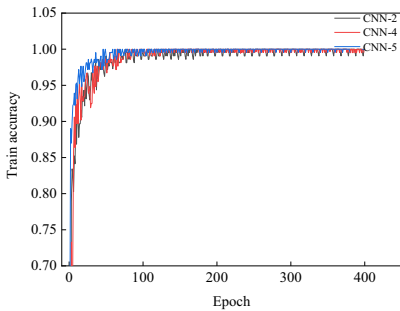
(d) the cross-entropy loss of CNN-1, CNN-3 and CNN-4



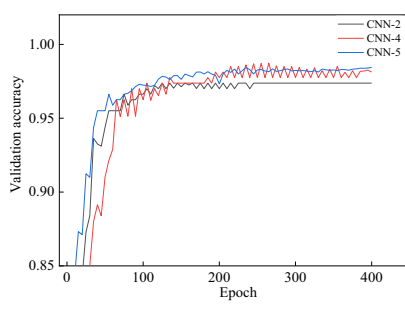
(e) the validation accuracy rate of CNN-1 and CNN-2



(f) the validation accuracy rate of CNN-1, CNN-3 and CNN-4



(g) the training accuracy rate of CNN-2, CNN-4 and CNN-5



(h) the validation accuracy rate of CNN-2, CNN-4 and CNN-5

Fig. 14 The training accuracy, validation accuracy, cross entropy of each model. **a** the training accuracy rate of CNN-1 and CNN-2, **b** the cross-entropy loss of CNN-1 and CNN-2, **c** the training accuracy rate of CNN-1, CNN-3 and CNN-4, **d** the cross-entropy loss of CNN-1, CNN-3 and CNN-4, **e** the validation accuracy rate of CNN-1 and CNN-2, **f** the validation accuracy rate of CNN-1, CNN-3 and CNN-4, **g** the training accuracy rate of CNN-2, CNN-4 and CNN-5, **h** the validation accuracy rate of CNN-2, CNN-4 and CNN-5

than the ReLU function in weld defect image recognition. Figures 14c, d, f show the training accuracy, cross-entropy loss and verification accuracy changes of convolutional neural network models using different pooling methods. The CNN-4 model achieved higher recognition rates than the CNN-1 model of the mean pooling method and the CNN-2 model of the max pooling method. It can be verified that the improved pooling method can be used in defects and weld transition areas to increase the defect feature domain, improve accuracy in feature selection, and the extracted features can better describe the image, which is conducive to the improvement of the defect recognition rate. Figures 14g, h are the images of the training and verification accuracy changes using the improved activation function and adaptive pooling method. It can be seen from the figure that using both methods can obtain very good results and performance in weld flaw detection image recognition at the same time.

In order to further verify the validity and reliability of the model in the identification of weld inspection images, another part of the weld inspection images were taken for testing. In the test results, 2 pieces of each of the 5 defect categories were randomly selected and numbered. The recognition results are shown in Table 2. The classification results indicate shows the probability of the image to be identified as non-defective, non-welded, unfused, stomatal, and cracked after calculation by the softmax layer. It can be seen that the overall effect of different convolutional neural network models recognizes non-defective, stomatal, and non-welded is good, while the recognition effect of unwelded and crack defects is relatively poor. The reason may be that the amount of training data for the two defects is too small, which causes the model to insufficiently learn the feature of defects. For the same defect type, the improved convolutional neural network model has a better recognition rate, which shows that the method in this paper has more advantages in feature extraction. Through the overall test sample recognition analysis, the method proposed in this paper can effectively identify 5 types of weld inspection images, and the overall recognition accuracy is 98.13% for CNN-5. It can be seen that the method described in this article can improve the recognition rate by 1.5% compared with the traditional convolutional neural network, and has a higher

Table 2 Partial sample recognition results

Experiment	Test sample	CNN-1	CNN-2	CNN-3	CNN-4	CNN-5
1	FL	99.9%	99.9%	99.7%	99.9%	99.9%
2	FL	99.9%	99.9%	99.8%	99.9%	99.9%
3	GP	99.9%	99.9%	99.9%	99.9%	99.9%
4	GP	99.9%	99.9%	99.9%	99.9%	99.9%
5	CK	55.9%	60.2%	51.1%	63.7%	70.2%
6	CK	49.9%	62.3%	52.7%	59.8%	68.4%
7	LOF	71.3%	85.8%	56.1%	86.2%	88.5%
8	LOF	82.4%	87.7%	79.6%	87.9%	88.0%
9	LOP	96.2%	96.8%	93.6%	97.7%	97.9%
10	LOP	90.7%	92.5	91.8%	97.3%	97.4%

Table 3 Recognition effect comparison

algorithm	Recognition accuracy
Literature [18]	93.5%
Literature [7]	96.77%
Literature [26]	97.0%
Algorithm in the text	98.13%

recognition rate. It is fully expected to achieve accurate subdivision and recognition of various defects in the weld image.

4.4 Compared with other methods

Through the verification and analysis of weld flaw detection images, the method used in this article has a good recognition rate in defect recognition. In order to further illustrate the superiority of this method, the improved method is compared with other model algorithms, and the comparative analysis is shown in Table 3.

Compared with other methods, the method adopted in this paper has better performance in defect recognition accuracy, generalization ability and robustness.

5 Conclusion

1. The image processing method can effectively segment the weld and defects, and the defect location in the weld image can be located.
2. The ELU activation function is used in the construction of the CNN model, which makes the model more robust during training, and increases the convergence speed through good network sparsity and a smaller mean value of output.
3. An improved pooling method based on grayscale feature adaptation can dynamically adjust the characteristics of weld images, reduce the impact of image noise on the training process, increase the extraction range of weld defect features, and have dynamic adaptability to certain pooling domains with different feature distributions.
4. The CNN model constructed by the ELU activation function and the improved pooling method based on adaptive grayscale features provided in this article can be used in the field of automatic detection of weld images, and can significantly improve the accuracy of defect recognition in the weld image, the overall recognition rate can reach 98.13%. We can see that the method described in this article can improve the recognition rate by 1.5% compared with the traditional convolutional neural network. It is completely predictable to realize the accurate subdivision and identification of various defects in the weld image, and the method is universal and can be extended to other fields.

Acknowledgments This work is funded by the National Natural Science Foundation of China (No. 01020607).

Declarations

Ethical approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

1. Ali A, Zhu Y, Zakarya M (2021) A data aggregation based approach to exploit dynamic spatio-temporal correlations for citywide crowd flows prediction in fog computing [J]. *Multimed Tools Appl* 2:4–37
2. Bharodiya AK, Gonsai AM (2019) An improved edge detection algorithm for X-Ray images based on the statistical range. *Heliyon* 5:27–36
3. Ding F, An-de H, Jian-kang H et al (2020) X-ray image defect recognition method for pipe weld based on improved convolutional neural network [J]. *Transactions of the China Weld Institution* 41(01):7–11 +97
4. Federer C, Xu H, Fyshe A et al (2020) Improved object recognition using neural networks trained to mimic the brain's statistical properties - ScienceDirect [J]. *Neural Netw* 131:103–114
5. Gao L, Liao YL, Zhu BH et al (2007) Adaptive thresholds edge detection for image based on segmentation. *Journal of Transport Information and Safety* 25(5):73–76
6. Gorban AN, Mirkes EM, Tugin IY (2020) How deep should be the depth of convolutional neural networks: a backyard dog case study [J]. *Cogn Comput* 12(1):388–397
7. Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J, Chen T (2018) Recent advances in convolutional neural networks. *Comput Sci* 77:354–377
8. He K, Zhang X, Ren S et al (2014) Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 37(9):1904–1920
9. Lecun Y, Bengio Y (2015) Hinton G. *Deep Learning* 521(7553):436–445
10. LeCun Y, Bottou L, Bengio Y et al (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
11. Li T, Meng Z, Ni B, Shen J, Wang M (2016) Robust geometric ℓ_p -norm feature pooling for image classification and action recognition. *Image Vis Comput* 55(2):64–76
12. Li Y, Xie W, Li H (2017) Hyperspectral image reconstruction by deep convolutional neural network for classification. *Pattern Recogn* 63:371–383
13. Lin JY, Ke HR, Chi BC et al (2007) Designing a classifier by a layered multi-population genetic programming approach. *Pattern Recogn* 40(8):2211–2225
14. McIlhagga W (2010) The canny edge detector revisited. *Int J Comput Vis* 91(3):251–261
15. Mery D, Riffo V, Zscherpel U, Mondragón G, Lillo I, Zuccar I, Lobel H, Carrasco M (2015) GDXray: the database of X-ray images for nondestructive testing. *J Nondestruct Eval* 34(4):1–12
16. O'Shea TJ, Hoydis J (2017) An introduction to deep learning for the physical layer. *IEEE Transactions on Cognitive Communications & Networking* 3(4):563–575
17. Sakamoto M, Hiasa Y, Otake Y, Takao M, Suzuki Y, Sugano N, Sato Y (2020) Bayesian segmentation of hip and thigh muscles in metal artifact-contaminated CT using convolutional neural network-enhanced normalized metal artifact reduction [J]. *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology* 92(3):335–344
18. SHAO W, LIU X, WU Z. (2019) A robust weld seam detection method based on particle filter for laser welding by using a passive vision sensor [J]. *Int J Adv Manuf Technol* 4:1–10
19. Sun Y, Song HH, Bai P et al (2004) Real-time automatic detection of weld defects in X-ray images. *Transactions of the China Welding Institution* 25(2):115–118
20. Wang F, Zhu JJ (2017) Research of machine vision for welding defect in complex environment. *Welding Technology* 46(5):127–133
21. Wei Y, Xia W, Lin M et al (2015) HCP: a flexible CNN framework for multi-label image classification. *IEEE Trans Softw Eng* 38(9):1901–1907
22. Wu X, He R, Sun ZN et al (2017) A light CNN for deep face representation with Noisy labels. *IEEE Transactions on Information Forensics & Security* 14(8):1–13
23. Yang Z, Kuang N, Fan L et al (2018) Review of image classification algorithms based on convolutional neural networks. *J Signal Process* 34(12):84–99
24. Yang M, Huang F, Lv X (2019) A feature learning approach for face recognition with robustness to noisy label based on top-N prediction. *Neurocomputing* 330(22):48–55

25. Yang L, Chen W, Liu W, Zha B, Zhu L (2020) Random noise attenuation based on residual convolutional neural network in seismic datasets [J]. *IEEE Access* 8:30271–30286
26. Yong-wei YU, Liu-qing DU, Zhe Y et al (2016) Dynamic detection of casting defects radiographic image based on deep learning feature. *Transactions of the Chinese Society for Agricultural Machinery* 47(7):407–412
27. Zhang ZT, Cheng DD, Wang JJ et al (2019) Quantum based subgraph convolutional neural networks. *Pattern Recogn* 88:38–49
28. Zhang Z, Ding S, Jia W (2019) A hybrid optimization algorithm based on cuckoo search and differential evolution for solving constrained engineering problems. *Eng Appl Artif Intell* 85:254–268
29. Zhang Z, Ding S, Sun Y (2020) A support vector regression model hybridized with chaotic krill herd algorithm and empirical mode decomposition for regression task. *Neurocomputing* 410:185–201
30. Zhou FY, Jin LP, Dong J (2017) Review of convolutional neural network. *Chinese Journal of Computers* 40(6):1229–1251

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.