

Oil Pipeline Weld Defect Identification System Based on Convolutional Neural Network

Jiaze Shang^{1*}, Weipeng An¹, Yu Liu¹, Bang Han¹, Yaodan Guo¹

¹ School of computer science and technology, Henan Polytechnic University
Jiaozuo Henan, 454000, China
[e-mail: shangjiazecn@163.com]

*Corresponding author: Jiaze Shang

*Received May 5, 2019; revised October 9, 2019; accepted November 24, 2019;
published March 31, 2020*

Abstract

The automatic identification and classification of image-based weld defects is a difficult task due to the complex texture of the X-ray images of the weld defect. Several depth learning methods for automatically identifying welds were proposed and tested. In this work, four different depth convolutional neural networks were evaluated and compared on the 1631 image set. The concavity, undercut, bar defects, circular defects, unfused defects and incomplete penetration in the weld image 6 different types of defects are classified. Another contribution of this paper is to train a CNN model "RayNet" for the dataset from scratch. In the experiment part, the parameters of convolution operation are compared and analyzed, in which the experimental part performs a comparative analysis of various parameters in the convolution operation, compares the size of the input image, gives the classification results for each defect, and finally shows the partial feature map during feature extraction with the classification accuracy reaching 96.5%, which is 6.6% higher than the classification accuracy of other existing fine-tuned models, and even improves the classification accuracy compared with the traditional image processing methods, and also proves that the model trained from scratch also has a good performance on small-scale data sets. Our proposed method can assist the evaluators in classifying pipeline welding defects.

Keywords: pipeline x-ray welding image; automatic identification; feature extraction; convolution neural network; convolution operation

1. Introduction

Welding, as the basic process of connecting components, has been widely used in the industrial production of railway bridges, aerospace, petroleum equipment manufacturing [1]. In the manufacturing process of welded special steel pipes for oil transmission, various welding defects such as concavity, undercut, bar defects, circular defects, unfused defects and incomplete penetration are unavoidable due to the constraints of production environment and production process. The description of pipeline welding defects is shown in Fig. 1 [2]. Therefore, it is necessary to detect and identify the welding defects of oil pipelines. X-ray nondestructive testing (NDT) is a testing method which uses X-ray to penetrate metal materials, and because of the different absorption and scattering effects of materials on the rays, the film has different photosensitivity and forms images with different blackness on the negative film, according to which the defect inside materials can be judged [3][4]. However, due to the features of X-ray weld film itself, the extracted X-ray image has many problems, such as noise, low contrast, large background fluctuations, complex weld image texture, and blurred weld edges, which make the automatic extraction of welding defects still a major problem [5]. And even if the defects are extracted, because the classification standard of welding defects has not been quantified, the automatic identification and classification of defects are also major problems. The traditional detection method relies on manual evaluation, This evaluation method is affected by the equipment, environment, working time and personal experience, and for the evaluators, not only the workload is large, but also the efficiency is low. It is also easy to cause inconsistent evaluation results due to the difference of personal experience, thus resulting in missing judgment and misjudgment of defects.

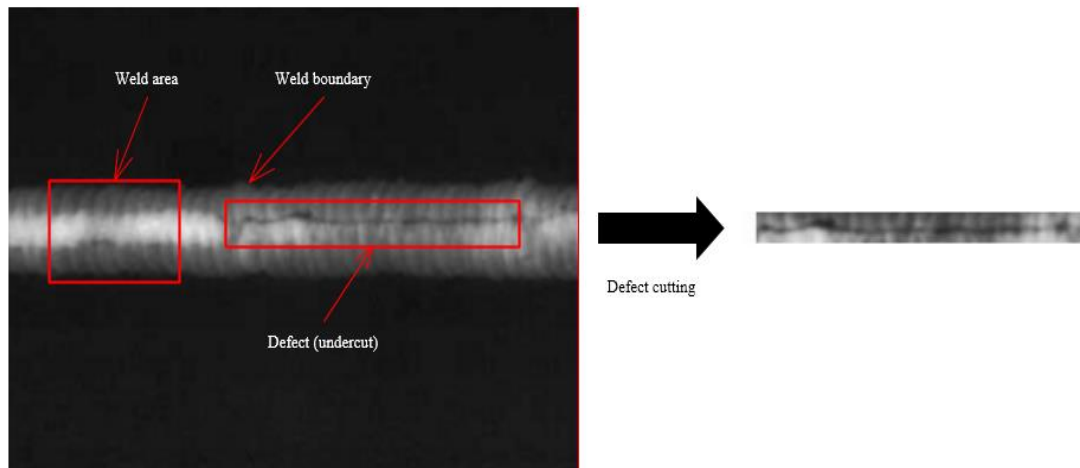


Fig. 1. Description of pipeline welding defects

At present, the detection method has received attention from many countries and enterprises that digital image processing technology is used to process and analyze the pipeline welding images, and then determine whether there are defects in the pipeline. The ultimate goal of image processing is to extract regions of interest and analyze them accordingly [6]. In this process, we should overcome the shortcomings of image noise, a large amount of useless information and low gray contrast. There are different algorithms and processing methods for image processing technology in various industries, but the basic method is to perform image

enhancement after pre-processing for noise reduction, and then determine the threshold value of binary segmentation, use this threshold value to binarize the image, and separate the target region from other parts [7][8][9][10]. The target area is marked by computer and the characteristic parameters, such as area, perimeter, circularity, and gray level that can describe it, are extracted. Finally, the whole target is analyzed or identified by using the characteristic parameters to achieve the desired purpose. The traditional identification methods are all identification methods that rely on feature engineering, while it is very difficult to extract features accurately and effectively, and it is difficult to determine many feature types such as set features and texture features [11][12][13]. This leads to that the identification accuracy cannot be further improved, and artificial feature engineering needs artificial determination of features, which requires a lot of manual work, thus significantly reducing the degree of automation of defect identification. This bottleneck faced puts forward requirements for us to break through the traditional methods and find new defect identification methods.

The essence of deep learning is to learn the deep essential features of the given data from the large-scale data through deep learning network structure [14][15][16], so as to improve the accuracy of classification and identification. LeCun et al. defines the basic structure of modern convolution neural network, in which alternating convolution layers and pooling layers are considered to be effective in extracting translation invariant features of input images [17]. Other applications based on convolutional neural networks, including face identification and gesture identification, have also been developed. Convolution neural networks require a large data set in image classification. However, it is difficult to obtain X-ray photographs of large defects in the X-ray field. In reference [18], the author trained CNN models with less data to have obtained higher accuracy rates. Due to the emergence of high-performance GPUs, the CNN model has been developed faster and its training time has been shortened more than before.

The rest of the paper are as follows: Section 2 introduces the previous applications of digital image processing techniques in pipeline welding defect identification and analyzes these methods in depth. Section 3 describes the application of deep learning methods in welding defect identification. Section 4 gives the experimental analysis and comparison. Section 5 draws the final conclusions.

2. Related Work

In this Section, we introduced the latest research background of traditional pipeline welding defect identification, and analyzed its advantages and disadvantages, which enlightens the application of convolution neural network in welding defect identification. As early as 1992, Lei et al. proposed an unsupervised model-based image segmentation technique for X-ray CT [19], which utilizes a finite mixture distribution and an underlying Gaussian random field as a random image model to classify images by a Bayesian classifier. Abdelhak et al. first used homomorphic filter to preprocess the images [20]. Then, a global threshold method was used to extract the weld on the whole image. When the welding area was roughly determined, the local threshold method was used as the final threshold to extract the welding defects. This method can extract some relatively obvious defects effectively to a certain extent, but the effect of extracting defects of complex shape is not obvious. Malarvel et al. proposed an improved anisotropic diffusion model based on X-ray pipe weld images [21], which considers the adaptive threshold parameters in the local gray probability value and diffusion coefficient function to achieve the meaning of adjusting the low edge gradient in the noise image feature space. The performance of the model was evaluated using indices such as mean square error, signal-to-noise

ratio, and entropy, and the reliability of the model was verified. Nafaa et al. proposed a computer-aided diagnosis software for welding defects and proposed an unsupervised classifier based on a finite mixture model of Multivariate Generalized Gaussian Distribution (MGGD) [22].

The parameters of the hybrid model based on Non-zero Mean MGGD were estimated by using an expectation-maximization (EM) algorithm, and the classification accuracy was over 96%. Shao et al. used a defect segmentation algorithm with low threshold to segment the potential defects of the image [23], and used an improved Hough transform to track the center of gravity of the potential defects in the image sequence and removed the potential defects that could not be tracked as pseudo-defects. This method has the advantages of fast speed, strong adaptability, good stability and so on. However, these are traditional image processing methods, which need to go through the traditional denoising, image enhancement, image segmentation and other operations to get the defects in the weld. After extraction, it still needs to be judged manually, which is quite backward for the current industrial level.

Although these methods have achieved high classification accuracy, the generalization performance of these methods is very low, different results may be obtained for different welding defects, and some small defects in the images may be missed after various image processing methods are used, and these shortcomings are unavoidable in these approaches.

At present, the convolution neural network has been applied in many image analysis, which can avoid these shortcomings. Ciocca et al. used food images in a variety of CNN models and compared the classification accuracy of these models [24]. C. Spampinato et al. evaluated the age of bones with convolutional neural networks and obtained high assessment accuracy [18]. The experiment was operated on 1391 photographs of the palm, which was unprecedented in bone age assessment. At the same time, Al-antari et al. designed a breast cancer computer diagnosis system for convolution training on 410 mammograms [25], and the detection accuracy of this system reached 98.96%, which was superior to the latest traditional deep learning methods. The applications in medical X-ray images also include the use of multiple convolution neural networks to classify chest images including 14 diseases and pathologies, and the classification accuracy was also 81.6%, which was unprecedented in medical image classification [26]. The application of X-ray image in deep learning is also reflected in the detection of voids in crystal chips, and Van Veenhuizen trained on the convolutional neural network with 219 crystal X-ray images and obtained a high identification rate [27], which proved that small data sets can also perform well in convolutional neural networks. Jiwang Jiang et al. obtained cross-sectional images of bituminous hybrid materials using X-ray Computed Tomography (CT) method and trained them in a CNN network model [28], and also obtained good classification results. As far as we know, convolution neural network has not been applied to identify welding defects in oil pipelines. However, machine learning methods have been applied in defect recognition and Valavanis et al. used local threshold method to segment the image [29], and then used the existing multi-classification support vector machine and neural network classifier to classify the defects, and the experiments showed that the computation time can be significantly reduced without affecting the overall accuracy. Zapata et al. designed an automatic detection, identification and classification system of welding defects based on an artificial neural network and an adaptive network-based fuzzy inference system [30]. The system can identify four main types of welding defects. The accuracy of the neural network prediction is 78.9%, and it can be seen that the accuracy is relatively low. Some real-time detection methods have also been applied to the detection of welding defects. Zou et al. proposed a real-time detection method for welding defects based on Kalman filter [31]. Kalman filter uses to detect the trajectory continuity of defects in image

sequences to identify real defects, but this method has great limitations and cannot identify the discontinuous circular defects.

At present, the more popular methods in deep learning are: using the ready-made to train directly on the target dataset, improving the existing model, and training it. For example, Pranata et al. trained computed tomography images of calcaneal fractures using two readily available convolutional neural network structures and finally obtained a classification accuracy of 98% [32]. Nasr-Esfahani et al. trained with the CNN model in X-ray angiography images for the purpose of detecting coronary heart disease [33].

In this paper, our contributions include: first training on pipeline weld images using convolutional neural networks and identification and classification of defects in welds. Firstly, the dataset is trained in the existing CNN model. Secondly, it is fine-tuned based on the existing model, and the classification accuracy is improved by training the data in the fine-tuned model. Finally, we train a unique model for the welding dataset from scratch, which is able to automatically locate the defect positions on welding images and enable accurate identification and classification. This is unprecedented.

3. Proposed Method

In this Section, we proposed three different methods for detecting welding defects in oil pipelines: (1) using transfer learning methods, that is, classifying images using existing CNN models, which does not require any model adaptation; (2) fine-tuning the existing CNN model to adapt the model to our pre-trained dataset, and the role of fine-tuning is to update the weights in the network, which has better results compared to the first one; and (3) starting training a CNN model against the heavy head of our existing dataset. The three models have similar structures and consist of two consecutive networks: (1) a convolution layer that performs feature extraction on the input data, and the number of layers of the convolution layer can be varied according to the actual situation; and (2) a fully connected layer that acts as a "classifier" in the whole convolutional neural network. Finally, an output layer consisting of logical functions or normalized exponential functions (Softmax function) was used to output taxonomic labels. The current mainstream, classical CNN structural model is shown in Fig. 2.

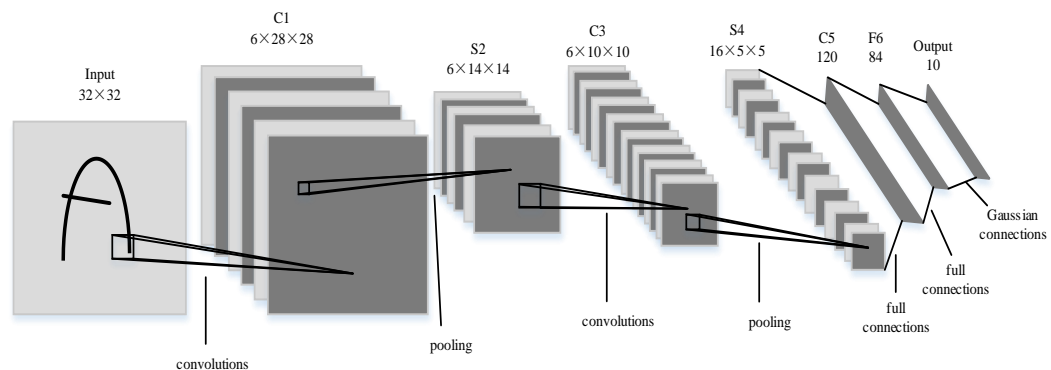


Fig. 2. A classical convolutional neural network model

3.1 Transfer Learning for Feature Extraction

Deep convolutional activation is a transfer learning method, and the general practice is to extract the output of the final layer of convolutional or penultimate fully connected layers as CNN features in existing CNNs and then classifies them directly using classifiers such as

SVM, Bayesian or Softmax. At present, the CNN models with high accuracy include LeNet, OverFeatNet, AlexNet, Google LeNet, Oxford Net, ImageNet and VGGNet. These models are trained with large datasets and have the ability to extract shallow basic features and deep abstract features. The transfer learning method is very effective, but if it is less similar, for example, the pre-trained dataset used is a natural landscape picture but is used for face recognition, the result is not so good, because the feature extraction of the face is different from the feature extraction of the natural scenery, and the convolution kernel learning does not have the same recognition pattern, so the parameters of the corresponding training are also different. Obviously, the X-ray images of the data set used in this paper are significantly different from the data set used in the pre-trained CNN model. Among the existing CNNs models, we consider three different pre-trained CNNs models: OverfeatNet, VGGNet, and GoogleNet.

For these three network models, we need to adjust the size of the image to make it compatible with the input layer of each CNNs model. The input image sizes of the three models are 221×221 , 224×224 , 299×299 , respectively. The input image sizes corresponding to the above three models are 221×221 , 224×224 , 299×299 , respectively. Due to the different sizes of data images obtained after sampling, all images need to be up-sampled or down-sampled to meet the requirements of each model input layer.

These three kinds of pre-trained networks have powerful functions in general visual tasks, especially in image classification tasks. Weights of models in the convolutional layer and fully connected layer are applicable to shape, texture and complex features of all images. The fully connected layer of the above three models can reach 2048 neurons (Overfeat), 4096 neurons (VGGNet) and 1000 neurons (GoogleNet). In this transfer learning model, we chose Support Vector Machine (SVM) as the classifier to classify the neurons processed through the fully connected layer, and then output seven classification results.

3.2 Fine-tuning the CNN model

Fine-tuning CNN is also one of the applications in deep learning methods. Fine-tuning this method can save a lot of computational resources and computation time, improve the computational efficiency and even accuracy, improve the generalization ability of the model, effectively prevent over-fitting, and achieve this by modifying the weights in the convolution layer and fully connected layer or modifying the number of output classes for different data sets. Therefore, OverfeatNet, VGGNet and GoogLeNet are fine-tuned for our pipeline welding defect dataset, in which the final classification number is adjusted from 1000 to 6 (six types of defects are included in the pipeline welding defect dataset), and the pipeline welding defect dataset is back-propagated through the fine-tuned model above, so as to obtain new weights and learning rates.

3.3 RayNet structure

Our last CNN model, named RayNet, was trained from scratch in the existing data set of welding defects in oil pipelines. The advantage of training a completely new CNN model over a fine-tuned one is that the entire CNN network architecture can be tuned according to the existing dataset: optimizing the grayscale image input size, reducing the number of convolution layers, and letting the network learn specific filters instead of adapting to generic filters. As shown in [Fig. 3](#), RayNet consists of a convolution layer, a pooling layer, a ReLU function not shown, a fully connected layer, and finally a classification layer.

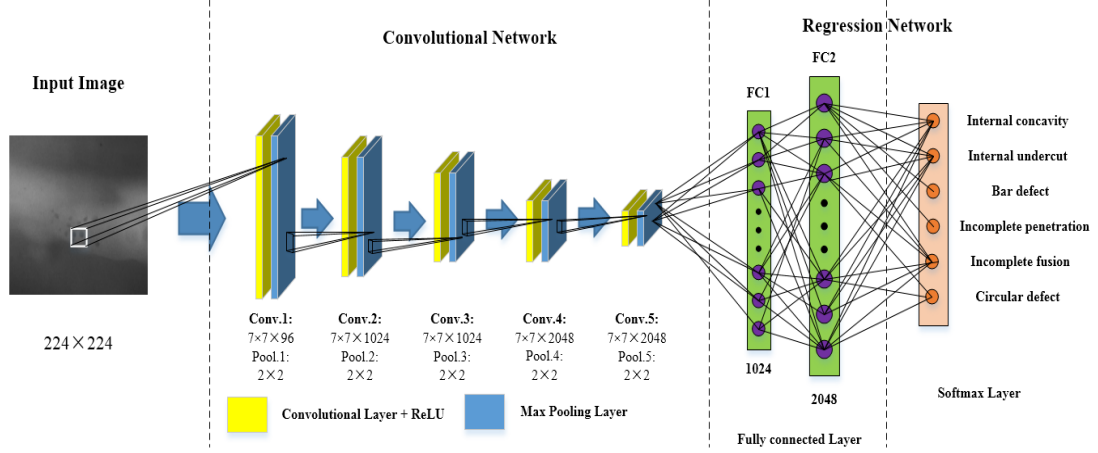


Fig. 3. Detailed model of different stages of RayNet: consisting of five convolution layers (pooling layer, ReLU function), two fully connected networks with neuron numbers of 1024 and 2048, respectively, and finally a Softmax classification layer.

The input image needs to be preprocessed before the convolution layer, and the input image is scaled. We resized the input image and test the performance of RayNet with five different input image sizes. Finally, we chose the image resolution of 224×224 as the input layer of the network (experimentally verified in Section 4.4).

LeCun et al. constructed a relatively complete convolutional neural network structure, and added a pooling layer to filter the input features based on the original convolution layer [17]. In addition, the ReLU activation function is also widely used in modern CNN network structure. The function of the convolution layer is to extract features from the input data, which contains several convolution kernels, and each element of the convolution kernel corresponds to a weight coefficient and a deviation. For the input image (i, j) , the entire convolution formula is as follows:

$$T^{l+1}(i, j) = \sum_{k=1}^{K_l} \sum_{x=1}^f \sum_{y=1}^f [T_k^l(s_0 i + x, s_0 j + y) * w_k^{l+1}(x, y)] + b \quad (1)$$

$$(i, j) \in \{0, 1, \dots, L_{l+1}\} \quad L_{l+1} = \frac{L_l + 2p - f}{s_0} + 1 \quad (2)$$

In the formula, b is the amount of deviation, and T^l and T^{l+1} denote the input and output of the feature map of the $l+1$ layer, respectively. L_{l+1} is the size of the dimension of T^{l+1} . $T(i, j)$ is the pixel of the corresponding feature map (i, j) . K is the number of channels of the feature map. The number of channels of the image in our dataset is generally 3, so $K = 3$, and f , s_0 and p are the numbers of convolution layers, corresponding to the convolution kernel size, convolution steps, and number of fill layers, respectively.

After feature extraction in the convolution layer, the output feature map is transferred to the pooling layer for feature selection and information filtering. Its function is to replace the result of a single point in the feature map with the statistics of the feature map of its neighboring regions. We choose the maximum pooling operation when training the RayNet structure, whose expression is as follows:

$$x_{i+1}^{size} = \left(\frac{x_i^{size} - F_i + 2 \times Pad}{Stride} \right) \quad (3)$$

Where x_i^{size} represents the size of the input feature map, F_i represents the filter size, Pad represents the number of padding layers, and $Stride$ represents the pooling layer step size.

The ReLU activation function is to add nonlinear factors in the convolution layer and make the output of the convolution layer into a nonlinear mapping. Compared with other activation functions, the ReLU activation function can overcome the vanishing gradient problem, speed up training, perform well without supervised pre-training, and provide the sparse expression ability of the neural network. Its expression is as follows:

$$\sigma(z) = \max(0, z) \quad (4)$$

It is invariant when z is greater than or equal to 0; if it is less than 0, it is 0 after activation.

In order to enable RayNet to perform better, we need to test the goodness of the network from different aspects and enable the performance of the network to be better compared to existing CNNs, where these aspects considered include: (1) the number of convolution layers; (2) the number of feature maps in each convolution layer. Therefore, for our experiments, we will continuously test the effect of different numbers of convolution layers on the experimental results, and each convolution layer has 20 feature maps (experimentally verified in Section 4.4).

In our model, the convolution kernel used in the first convolution layer is 7×7 , the convolution kernel used in the second convolution layer is 5×5 , and the convolution kernel used in the remaining convolution layers is 3×3 . The fully connected layer after the whole convolution layer consists of 2048 neurons, and finally a Softmax regression function is added.

4. Experimental Results

In this section, we first introduce the experimental dataset, experimental settings, and then use different models to test and compare the dataset, and finally visualize the classification results.

4.1 Dataset

In this study, we utilized a private database of welding defects in oil pipeline to test and evaluate the various methods mentioned in the paper. The data set includes 1631 X-ray films with annotated welding defects, which include: concavity, undercut, bar defects, circular defects, unfused defects, and incomplete penetrated. **Fig. 4** presents an example diagrams of the six defects and shows different numbers of the six defects in **Table 1**. The original image size in the dataset is 256×256 , and we randomly cropped the image to make it meet the requirements of OverfeatNet, VGGNet, GoogleNet input layer. At the same time, we compared the input image size of RayNet in Section 4.

In order to train deep learning models, a large number of annotated datasets are required. The available image datasets pose a challenge for this training task, but the current mainstream approach is to perform data expansion and transfer learning on images. Data augmentation is a widely accepted process, which is used to increase the size of the training dataset with a faster convergence rate and avoid over-fitting of the model. In this study, the original images of oil pipeline welding defects were rotated at 45° angles (0° , 45° , 90° , 135° , 180° , 225° , 270°)

and 315°), and the original images were enhanced eight times. Therefore, we obtained a total of 13,048 images of welding defects in oil pipelines for training the various models proposed in the paper. Eighty percent of them were used to train models, 10 percent to validate models, and the remaining 10 percent to test models.

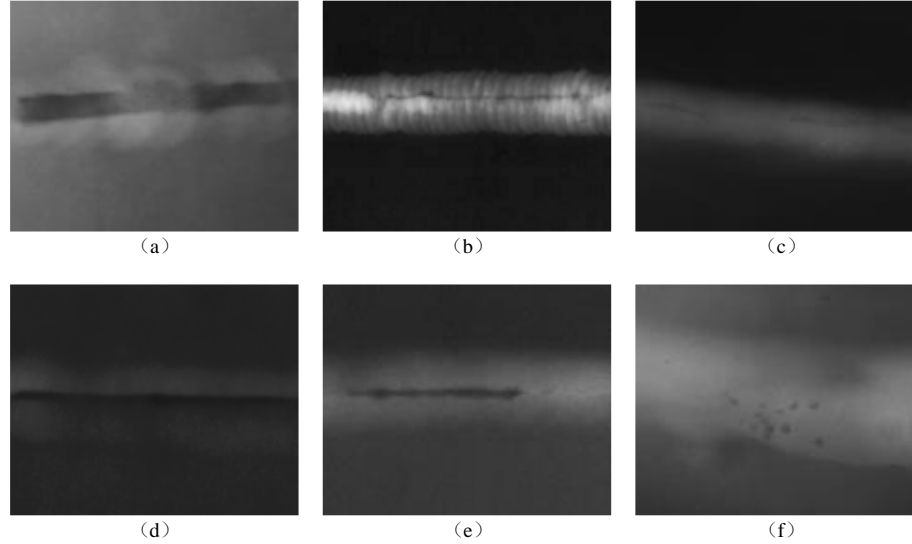


Fig. 4. Diagram of the six defects : (a) denotes concavity, (b) denotes undercut, (c) denotes bar defect, (d) denotes incomplete penetrated, (e) denotes unfused defect, and (f) denotes circular defect.

Table 1. Number of Different Defects in the Dataset

Defect Type	Concavity	Undercut	Bar defect	Incomplete penetrated	Unfused defect	Circular defect
Quantity	287	272	258	245	276	293

4.2 Experimental Setup

In training, the input image size after data expansion is adjusted to 224×224 . We used the stochastic gradient descent algorithm (SGD) to optimize the parameters of the network model. We used several groups of parameters to experiment with the proposed model. The experimental results are shown in **Table 2**, and it is concluded that the classification of group (A) parameters is the best. Therefore, we set batch size to 50 and trained 50 epochs. Learning rate was initially set to 0.001 and decreased to 0.0001 after 30 epochs, momentum factor was set to 0.9, and weight decay was set to 0.0001.

Table 2. Comparison of Various Experimental Parameters

Combination	A	B	C	D	E	
Learning Rate	0.001 and 0.0001	0.001 and 0.0001	0.005 and 0.0005	0.0001 and 0.00001	0.0005 and 0.00005	
Weight Decay	0.0001	0.0002	0.0001	0.0001	0.0001	
Batch Size	50	30	50	50	50	
Accuracy	96.5%	95.2	94.7%	92.1%	90.5%	
Iteration	Training	50	72	103	372	586
	Validation	16	23	29	46	113
Computation Time(Seconds)	1326	1469	1582	1769	2038	

For each stage, the generated training, validation and test data sets were subjected to four-fold cross-validation to ensure that each welding defect X-ray film was tested identically to prevent any errors and enable the network to obtain the best classification performance. All of these experiments were performed on computers with Intel (R) Core (TM) i7-8700K and 16GB RAM, equipped with NVIDIA TITAN Xp GPU, and the time of experimental training was 2840 seconds per epoch, while the test time was 1 – 2 seconds. All models were implemented in the TensorFlow environment.

4.3 Performance of pretrained convolutional neural networks

In this section, we introduced the results achieved by the pre-trained network as well as the fine-tuned network on the dataset. In the experiments, we compared the classification accuracy of different types of CNN's pre-trained using pre-trained networks for image classification and visual recognition applications, which include OverFeatNet, VGGNet, GoogleNet. As shown in Fig. 5, GoogleNet achieved the highest accuracy, followed by VGGNet, while OverFeatNet had the lowest classification accuracy of 71.8%.

In the fine-tuning test, we modified the final classifier to six classification results and compared it with the pre-trained network model experimentally, and the results are shown in Fig. 5. In order to increase the readability of the data, the accuracy rate in the map starts from 60%, and the results achieved by fine-tuning the network are predictable, because the parameters of the fine-tuned network structure will be adjusted according to different datasets, and at the same time the fine-tuned GoogleNet has the highest classification accuracy, which is improved by 14.7% on the basis of the existing model classification accuracy, because GoogleNet has a deeper network layer, and does not use the fully connected layer, etc.

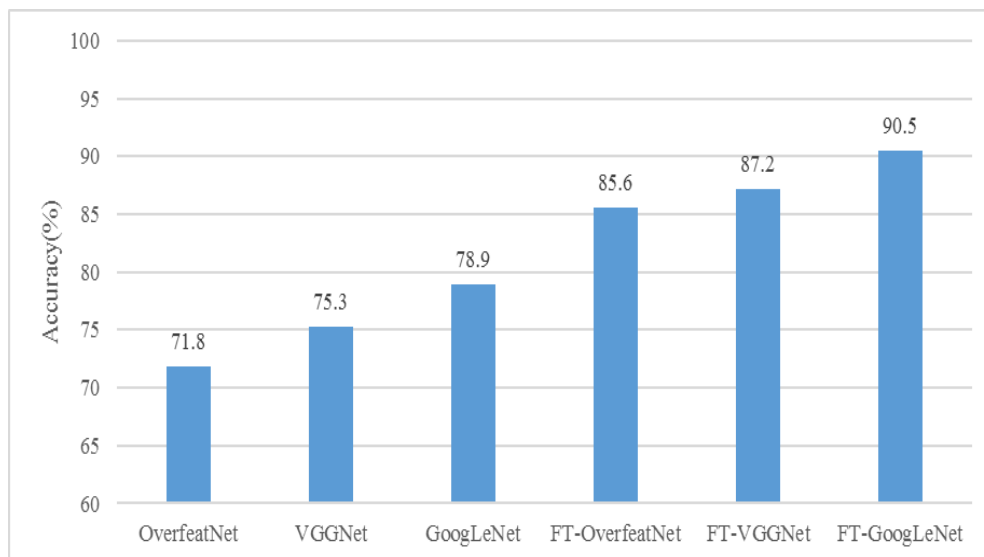


Fig. 5. Classification accuracy rates for different models and after fine tuning. FT-X "indicates the corresponding model after fine-tuning.

4.4 Performance of RayNet

In the experiment, we cut the original input images randomly, and the experimental results are shown in Fig. 6. when the image input layer is 224×224 , the classification accuracy reaches the highest 96.5%, and when the input layer is adjusted to 299×299 , the classification

accuracy reaches the lowest 92.7%. Therefore we take 224×224 as the input layer of RayNet based on the experimental results.

In the following experiments, the effect of network depth on classification accuracy was compared, including the number of convolutional layers, as well as the number of feature maps for each convolutional layer. We trained two to six convolutional layers, and the effect of different numbers of feature maps for each convolutional layer on the network performance, and the experimental results are shown in [Table 3](#). The experimental results show that when the number of convolution layers is 5 and the number of feature maps is 96, 1024, 1024, 2048, 2048 respectively, the network classification accuracy reaches the highest 96.5%, which shows that the network model can achieve good classification effect even if the number of convolution layers is much less than the number of existing models.

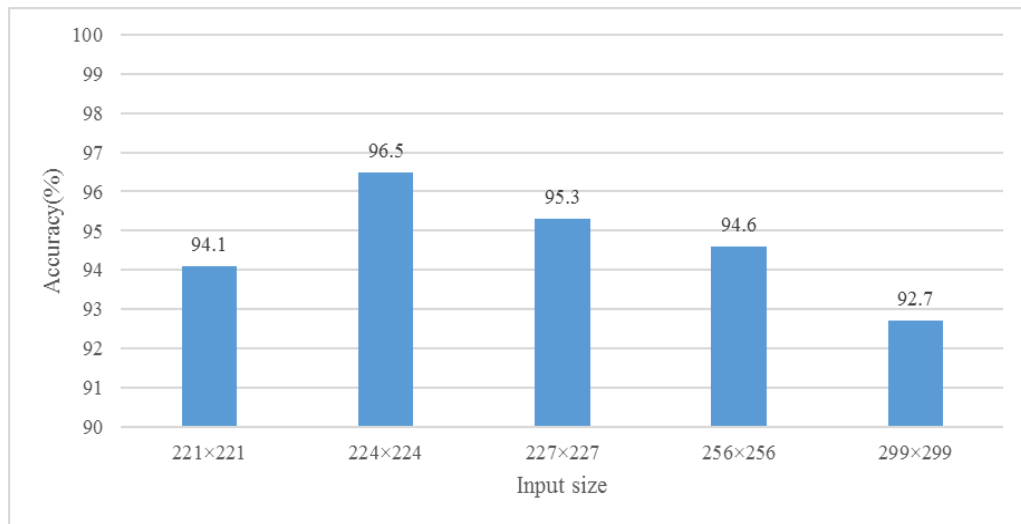


Fig. 6. Performance of accuracy rate in different input layers

Table 3. Performance of different convolutional layers and feature maps. "Conv." denotes the number of convolution layers, "map" denotes the number of feature maps for each convolution layer, and "accuracy (%)" denotes the classification accuracy corresponding to different structures.

Conv.	map	Accuracy(%)
2	96,1024	87.6
2	96,2048	88.2
3	96,1024,1024	88.9
3	96,1024,2048	89.1
3	96,2048,2048	89.2
4	96,1024,1024,1024	88.7
4	96,1024,1024,2048	89.7
4	96,1024,2048,2048	90.1

4	96,2048,2048,1024	90.8
5	96,1024,1024,1024,1024	93.5
5	96,1024,1024,1024,2048	95.7
5	96,1024,1024,2048,2048	96.5
5	96,1024,2048,2048,2048	95.9
5	96,1024,2048,2048,1024	96.2
6	96,1024,1024,1024,2048,2048	91.3
6	96,1024,2048,2048,2048,1024	90.5
6	96,2048,1024,2048,1024,2048	89.6

Finally, we compared the performance of different defects of oil pipeline weld (concavity, undercut, bar defect, circular defect, unfused defect and incomplete penetrated) in fine-tuning OverfeatNet, VGGNet, GoogLeNet as well as RayNet, and the experimental results are shown in **Table 4**. The experiments show that these four networks have more effect on different defects. It can be seen that the model has a better classification effect on circular defects as well as unfused defects. This is because the defects in our data set show more of these two types of defects. In the real pipeline welding operation, the performance is also more circular defects and unfused defects.

Table 4. Classification Accuracy of Different Defects in Different Networks (%)

Group	OverfeatNet	VGGNet	GoogLeNet	RayNet
Concavity	93.4	94.6	95.2	96.7
Undercut	92.5	93.8	94.5	95.3
Bar defect	92.8	94.3	94.9	96.1
Incomplete penetrated	91.2	92.5	93.1	93.8
Unfused defect	93.8	94.0	94.9	95.4
Circular defect	95.8	96.4	97.1	98.3

In order to further assess the misclassification results of RayNet model and GoogLeNet model in six defects, a confusion matrix is introduced to analyze the classification results of the two models. The experimental results are shown in **Fig. 7**.

As seen from **Fig. 7**, **Fig. (a)** is the confusion matrix obtained by the six kinds of defects in the RayNet model experiment. The blue part of the data on the diagonal line in the matrix means higher classification accuracy. In the undercut, there is 4% of the data is divided into bar defects. The appearance of this situation indicates that the two defects are similar, so it will cause misclassification. At the same time, this similarity in incomplete penetrated and unfused defect. However, the complete and correct classification can be achieved in circular defects, because the circular defect has a large amount of data, and the defects are more obvious, and the similarity with other defects is extremely low. **Fig. (b)** shows the performance of the six

defects on the GoogLeNet model, which is basically similar to the performance in Fig. (b), but performs worse in terms of classification accuracy.

As seen from Fig. 7, Fig. (a) is the confusion matrix obtained by the six kinds of defects in the RayNet model experiment. The blue part of the data on the diagonal line in the matrix means higher classification accuracy. In the undercut, there is 4% of the data is divided into bar defects. The appearance of this situation indicates that the two defects are similar, so it will cause misclassification. At the same time, this similarity in incomplete penetrated and unfused defect. However, the complete and correct classification can be achieved in circular defects, because the circular defect has a large amount of data, and the defects are more obvious, and the similarity with other defects is extremely low. Fig. (b) shows the performance of the six defects on the GoogLeNet model, which is basically similar to the performance in Fig. (a), but performs worse in terms of classification accuracy.

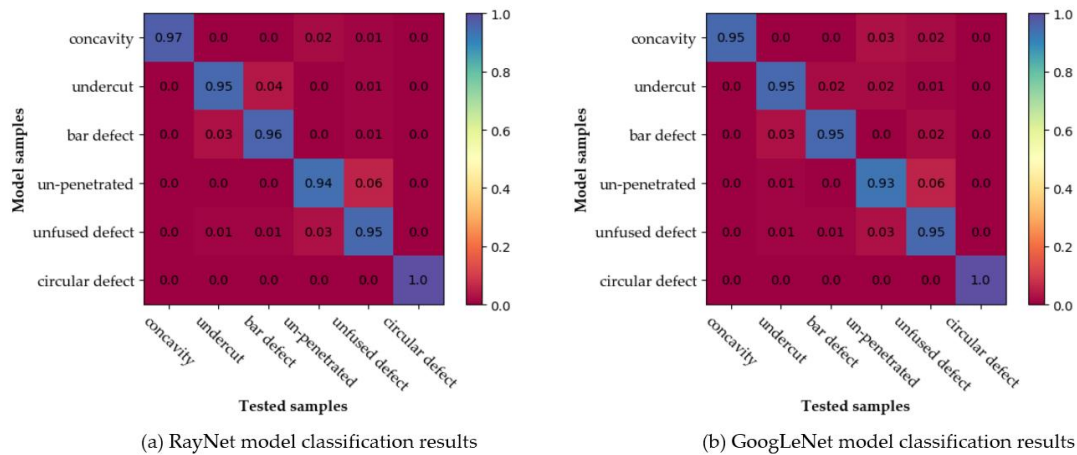


Fig. 7. The display of two model classification results in the confusion matrix (un-penetrated = incomplete penetrated)

Overall, RayNet is outperformed all the other CNN-based solutions, with a classification accuracy of 96.5% among the six defect types. Furthermore, the results of study also suggest that for weld defect X-ray images, many convolution layers — as in the case of GoogLeNet — are not strictly necessary. At the same time, in order to prove that the RayNet model is superior to other algorithm models, Table 5 shows the comparison of the average absolute error (MAE) between multiple algorithms in the literature and the RayNet model.

Table 5. Comparison of algorithms in the literature with RayNet in MAE

Method	Reading		Average
	1	2	
Malarvel et al. (2017)	2.03	1.95	1.99
Nacereddine et al. (2019)	1.92	1.87	1.90
Shao et al. (2012)	2.69	2.57	2.63
Al-Antari et al. (2018)	1.25	1.21	1.23
Guan et al. (2018)	1.34	1.28	1.31
Jiang et al. (2018)	0.98	0.85	0.915
RayNet	0.79	0.73	0.76

As seen in [Table 5](#), among the seven methods, the RayNet model has the lowest MAE. And the CNN model (Al-Antari et al. Guan et al., Jiang et al.) has a slightly lower MAE than the traditional image processing method (Malarvel et al., Nacereddine et al., Shao et al.). Therefore, we conclude that the accuracy of X-ray images processed with the CNN method is higher, and the specific CNN model for the weld-defect X-ray film of the pipeline is more suitable.

4.5 Feature Identification

We show the corresponding images of the RayNet model output at each layer separately in [Fig. 8](#), which shows the partial output results of the first two convolution layers due to the space problem. It can be seen from (b) that the features extracted from (a) the convolution operation on the input image (we just show the partially extracted features), (d) are the features extracted from the convolution operation on the basis of (b), and finally after a series of convolution, pooling as well as ReLU function operations we get the fully connected layer of (g).

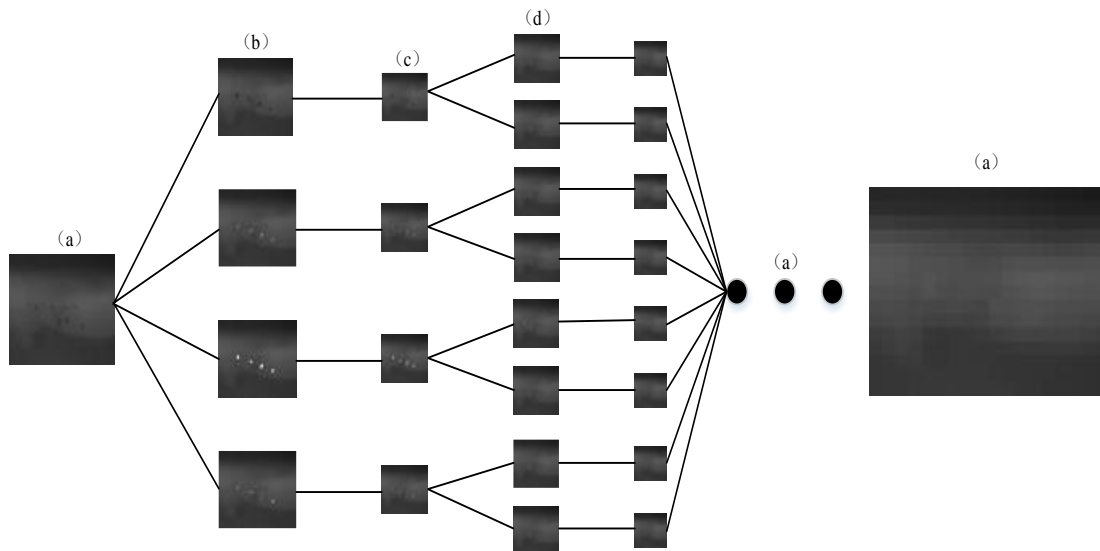


Fig. 8. Partial output results for each layer of RayNet network: (a) input image, (b) Conv. (a) output image, (b) output image, (c) Pool. 1 output image, (d) Conv. 2 output image, (e) Pool. 2 output image, (f) an output of the remain portion (shown differently for spatial reasons), (g) fully connected layer output result

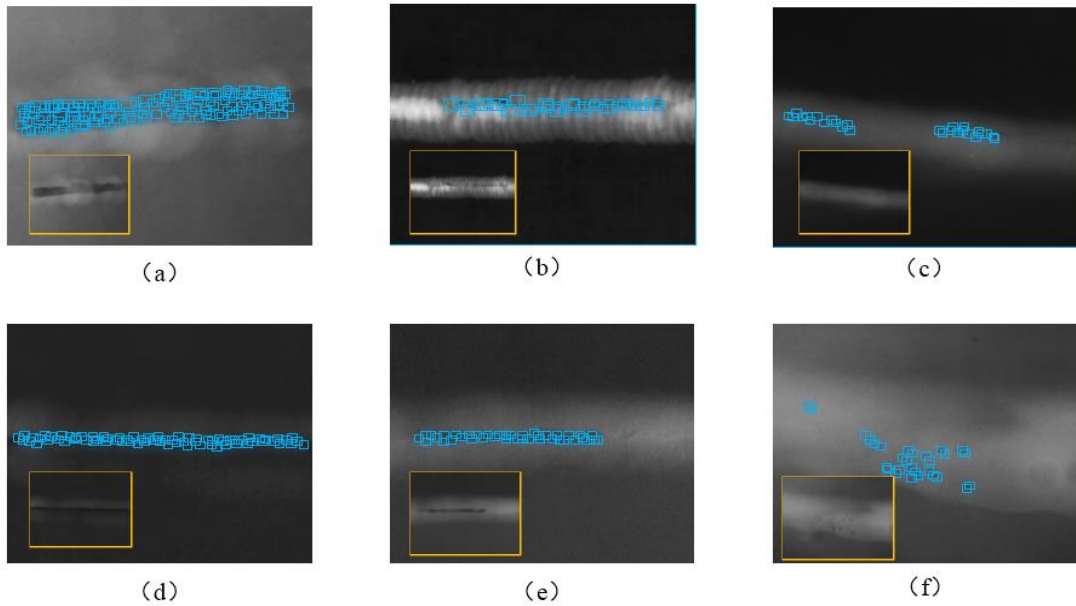


Fig. 9. Part of the feature points extracted from the RayNet model (the original image is the lower left corner of each picture.)

Fig. 9 shows the partial feature display extracted in the convolution operation engineering. We illustrate it with six different defects. The original image is shown in the lower left corner of each figure. It is very clear that the feature position extracted from the image is very accurate.

5. Conclusion

In this paper we have investigated the application of deep learning in X-ray images of oil pipeline welds. We have first tested the application of the existing pre-trained convolutional neural networks (OverfeatNet, VGGNet, GoogLeNet) in 13048 X-ray images. Fine-tuning the three models in the weld image, the accuracy rate is improved by 14.7%, which is better than the existing traditional image processing methods. The main contribution was to design and train a pipe weld X-ray proprietary convolutional neural network model (RayNet) from scratch, which consists of five convolutional layers and two fully connected layers. Experimental comparison of multiple sets of parameters yields a set of parameters that are more suitable for the model. At the same time, the influence of different input picture sizes on the model was tested, and 224×224 was finally determined as the input layer. The accuracy of classification and identification reached 96.5%. Compared with other pre-trained networks, the RayNet model has a smaller number of network layers and the performance demonstrated by the experiment is more intuitive. The experimental results are promising for the application of RayNet in the identification of pipeline welding defects, and this network model can assist the evaluators in the identification and classification of welding defects.

Although the model has achieved significant features, our study is limited by the lack of data volume and computational resources. In future work, we will acquire more pipeline weld images and collect more kinds of defects to improve the CNN model to improve the classification accuracy and performance when extracting features. In order to obtain better

performance, more effective detection methods are applied in automatic identification of pipeline welds.

References

- [1] Y. Crutzen, R. Boman, L. Papeleux and J. P. Ponthot, "Continuous roll forming including in-line welding and post-cut within an ALE formalism," *Finite Elements in Analysis & Design*, 143, 11-31, 2018. [Article \(CrossRef Link\)](#).
- [2] J. Shao, D. Dong, B. Chang and S. Han, "Automatic weld defect detection based on potential defect tracking in real-time radiographic image sequence," *NDT&E INT*, 46, 14-21, 2012.
- [3] G. Rooh, A. Khan, H. J. Kim, H. Park and S. Kim, "TlSr 2 Br 5: New intrinsic scintillator for X-ray and γ -ray detection," *OPT MATER*, 73, 523-526, 2017. [Article \(CrossRef Link\)](#).
- [4] T. Liu, J. Bao, J. Wang and Y. Zhang, "A Hybrid CNN-LSTM Algorithm for Online Defect Recognition of CO₂ Welding," *SENSORS-BASEL*, 18(12), 4369, 2018. [Article \(CrossRef Link\)](#).
- [5] C. C. Huang and H. N. Manh, "X-Ray Enhancement Based on Component Attenuation, Contrast Adjustment, and Image Fusion," *IEEE T IMAGE PROCESS*, 1, 2018.
- [6] J. Fan and J. Wang, "A Two-phase Fuzzy Clustering Algorithm Based on Neurodynamic Optimization with Its Application for PolSAR Image Segmentation," *IEEE T FUZZY SYST*, 26(1), 72-83, 2018. [Article \(CrossRef Link\)](#).
- [7] J. Miao, T. Huang, X. Zhou, Y. Wang and J. Liu, "Image segmentation based on an active contour model of partial image restoration with local cosine fitting energy," *INFORM SCIENCES*, 447, 52-71, 2018. [Article \(CrossRef Link\)](#).
- [8] Y. Li, G. Cao, Q. Yu and X. Li, "Active contours driven by non-local Gaussian distribution fitting energy for image segmentation," *APPL INTELL*, 48(12), 4855-4870, 2018. [Article \(CrossRef Link\)](#).
- [9] Hu, Y., Soltoggio, A., Lock, R., Carter, S., "A Fully Convolutional Two-Stream Fusion Network for Interactive Image Segmentation," *NEURAL NETWORKS*, 2018.
- [10] Ali, H., Rada, L., Badshah, N., "Image Segmentation for Intensity Inhomogeneity in Presence of High Noise," *IEEE Transactions on Image Processing*, 27(8), 3729-3738, 2018. [Article \(CrossRef Link\)](#).
- [11] Zhou, X., Gao, X., Wang, J., Yu, H., Wang, Z., Chi, Z., "Eye Tracking Data Guided Feature Selection for Image Classification," *PATTERN RECOGN*, 63, 56-70, 2017. [Article \(CrossRef Link\)](#).
- [12] Nizami, I.F., Majid, M., Khurshid, K., "New feature selection algorithms for no-reference image quality assessment," *APPL INTELL*, 48, 3482-3501, 2018. [Article \(CrossRef Link\)](#).
- [13] Zhou, W., Lu, Y., Qiu, W., Yang, Z., Wu, M., "Local gradient patterns (LGP): An effective local-statistical-feature extraction scheme for no-reference image quality assessment," *INFORM SCIENCES*, 397-398, 1-14, 2017. [Article \(CrossRef Link\)](#).
- [14] Minkyu L, Donghyun L, Hosung P, "Convolutional Neural Network based Audio Event Classification," *KSII Transactions on Internet and Information Systems*, 12(6), 2748-2760, January 6-12, 2018. [Article \(CrossRef Link\)](#).
- [15] Sheraz N, Yasir S, "Enhanced Network Intrusion Detection using Deep Convolutional Neural Networks," *KSII Transactions on Internet and Information Systems*, 12(6), 5159-5178, October 12-10, 2018. [Article \(CrossRef Link\)](#).
- [16] Giseok C, Seunghin L, Jongho N, "CNN-based Visual/Auditory Feature Fusion Method with Frame Selection for Classifying Video Events," *KSII Transactions on Internet and Information Systems*, 13(3), 1689-1701, January 13-3, 2019. [Article \(CrossRef Link\)](#).
- [17] Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., "Gradient-based learning applied to document recognition," *P IEEE*, 86(11), 2278-2324, 1998. [Article \(CrossRef Link\)](#).
- [18] Spampinato, C., Palazzo, S., Giordano, D., Aldinucci, M., Leonardi, R., "Deep Learning for Automated Skeletal Bone Age Assessment in X-Ray Images," *MED IMAGE ANAL*, 36, 41-51, 2017. [Article \(CrossRef Link\)](#).

- [19] T, L., W, S. Statistical approach to X-ray CT imaging and its applications in image analysis. II., "A new stochastic model-based image segmentation technique for X-ray CT image," *IEEE T MED IMAGING*, 11(1), 62-69, 1992. [Article \(CrossRef Link\)](#).
- [20] Mahmoudi, A., Regragui, F., "Fast segmentation method for defects detection in radiographic images of welds," in *Proc. of IEEE/ACS International Conference on Computer Systems & Applications*, 2009. [Article \(CrossRef Link\)](#).
- [21] Malarvel, M., Sethumadhavan, G., Bhagi, P.C.R., Kar, S., Saravanan, T., Krishnan, A., "Anisotropic diffusion based denoising on X-radiography images to detect weld defects," *DIGIT SIGNAL PROCESS*, 68, 112-126, 2017. [Article \(CrossRef Link\)](#).
- [22] Nacereddine, N., Goumeidane, A.B., Ziou, D., "Unsupervised weld defect classification in radiographic images using multivariate generalized Gaussian mixture model with exact computation of mean and shape parameters," *COMPUT IND*, 108, 132-149, 2019. [Article \(CrossRef Link\)](#).
- [23] Shao, J., Dong, D., Chang, B., Han, S., "Automatic weld defect detection based on potential defect tracking in real-time radiographic image sequence," *NDT&E INT*, 46, 14-21, 2012. [Article \(CrossRef Link\)](#).
- [24] Ciocca, G., Napoletano, P., Schettini, R., "Learning CNN-based Features for Retrieval of Food Images," in *Proc. of International Conference on Image Analysis & Processing*, 426-434, 2017. [Article \(CrossRef Link\)](#).
- [25] Al-Antari, M.A., Al-Masni, M.A., Choi, M.T., Han, S.M., Kim, T.S., "A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification," *INT J MED INFORM*, 117, 44-54, 2018. [Article \(CrossRef Link\)](#).
- [26] Guan, Q., Huang, Y., "Multi-label chest X-ray image classification via category-wise residual attention learning," *PATTERN RECOGN LETT*, 130, 259-266, 2020. [Article \(CrossRef Link\)](#).
- [27] Van Veenhuizen, M., "Void detection in solder bumps with deep learning," *MICROELECTRON RELIAB*, 88-90, 315-320, 2018. [Article \(CrossRef Link\)](#).
- [28] Jiang, J., Zhen, Z., Qiao, D., Ni, F., "Characterization and identification of asphalt mixtures based on Convolutional Neural Network methods using X-ray scanning images," *Construction & Building Materials*, 174, 72-80, 2018. [Article \(CrossRef Link\)](#).
- [29] Valavanis, I., Kosmopoulos, D., "Multiclass defect detection and classification in weld radiographic images using geometric and texture features," *EXPERT SYST APPL*, 37(12), 7606-7614, 2010. [Article \(CrossRef Link\)](#).
- [30] Zapata, J., Vilar, R., Ruiz, R., "Performance evaluation of an automatic inspection system of weld defects in radiographic images based on neuro-classifiers," *EXPERT SYST APPL*, 38(7), 8812-8824, 2011. [Article \(CrossRef Link\)](#).
- [31] Zou, Y., Dong, D., Chang, B., Ji, L., Pan, J., "Automatic weld defect detection method based on Kalman filtering for real-time radiographic inspection of spiral pipe," *NDT&E INT*, 72, 1-9, 2015. [Article \(CrossRef Link\)](#).
- [32] Pranata, Y.D., Wang, K.C., Wang, J.C., Idram, I., Lai, J.Y., Liu, J.W., Hsieh, I.H., "Deep learning and SURF for automated classification and detection of calcaneus fractures in CT images," *Comput Methods Programs Biomed*, 171, 27-37, 2019. [Article \(CrossRef Link\)](#).
- [33] Nasr-Esfahani, E., Samavi, S., Karimi, N., Soroushmehr, S.M., Ward, K., Jafari, M.H., Felfeliyan, B., Nallamothe, B., Najarian, K., "Vessel extraction in X-ray angiograms using deep learning," in *Proc. of 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 643-646, 2016. [Article \(CrossRef Link\)](#).



Jiaze Shang is currently studying at Henan Polytechnic University in Jiaozuo, Henan province, China. He is currently interested in deep learning, image processing and artificial intelligence.



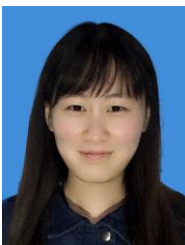
Weipeng An obtained his bachelor's degree and M.S. degree respectively at Henan Polytechnic University in Jiaozuo, Henan province, China. Since 1993, he has been a professor in the school of computer science and technology of Henan Polytechnic University. His research interests include computer measurement and control technology, deep learning and software development technology.



Yu Liu is currently pursuing a master degree at the School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China. Her current research interests include deep learning and pattern recognition.



Bang Han is currently pursuing a master degree at the School of Computer science and Technology in Henan Polytechnic University Wuhan, China. His current research interests include the information security.



Yaodan Guo is currently pursuing a master degree at the College of Computer Science and Technology in Henan Polytechnic University, Jiaozuo, China. Her current research interests include service computing and deep learning.