

文章编号:1007-5321(2019)06-0020-09

DOI:10.13190/j.jbupt.2019-097

# 基于 Faster R-CNN 模型 X-射线图像的焊接缺陷检测

郭文明, 刘凯, 渠慧帆

(北京邮电大学 软件学院, 北京 100876)

**摘要:** 为了实现 X-射线图片的焊接缺陷检测, 采用了基于目标检测领域的经典模型——Faster R-CNN 的目标检测方法. WDXI 数据集是从大量的 X-射线图像整理和分类构建获得的, 包括 7 种缺陷类型和无缺陷类型. 为了有效地提取焊接区域, 提出了一种根据平均灰度值和单位面积内平均对比度值的改进方法. 经过实验验证, 可采取自适应的直方图均衡化以及两次中值滤波的方法分别进行图像增强和降噪处理. 最终, 在焊接缺陷识别的多分类任务中, 训练模型在测试集上达到了预期效果, 不仅证明了 WDXI 数据集的研究价值, 还为实现焊接缺陷的自动识别和定位进行了实验性的尝试.

**关键词:** 焊接缺陷检测; 无损检测; X-射线图像; 目标检测; 卷积神经网络

中图分类号: TP183

文献标志码: A

## Welding Defect Detection of X-Ray Images Based on Faster R-CNN Model

GUO Wen-ming, LIU Kai, QU Hui-fan

(School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

**Abstract:** Based on faster region-based convolutional neural networks (R-CNN) model, a classical model in the field of object detection is used to achieve welding defect detection of X-ray images. A great number of X-ray images are collected and sorted out into the dataset, called WDXI, including no-defect type and 7 defect types. Firstly, an improved method can be used to extract the welding area effectively according to the average gray value and the average contrast value per unit area. The adaptive histogram equalization is used for image enhancement and double median blur is used for noise reduction after experimental comparison. Finally, the testing on the pre-trained model is expected and acceptable in the multi-classification problem of welding defects recognition, and not only proves the research value of WDXI, but also contributes towards making an experiment attempt for improving automatic classification and localization of welding defects combined with Faster R-CNN model.

**Key words:** welding defects detection; non-destructive testing; X-ray image; object detection; convolution neural networks

Automatic welding defect detection is a hot research direction in the field of non-destructive testing (NDT). People try different methods to recognize the

existing welding defects based on the texture features of the image or other information obtained from image. Different from the traditional NDT process, deep learn-

收稿日期: 2019-05-31

基金项目: 国家重大科学仪器设备专项项目(2013YQ240803)

作者简介: 郭文明(1967—), 男, 副教授, 硕士生导师, E-mail: guowenming@163.com.

ing also provides a solution for welding defect detection. It will put a large number of welding images into the neural network, relying on deep learning to recognize defects and defect types, and has a good performance on recognition.

There is a try to rely on deep learning for recognizing welding defects. Based on the Faster R-CNN model, a classifier is trained for welding defect recognition and localization to solve a multi-classification problem. And an improved method can be used to extract the welding area from original images.

## 1 Introduction

NDT is a technique which could recognize defects without damaging the detected objects by means of radiographic or ultrasound technology, which is widely used in many fields. In the welding processing, defects are inevitable. And these defects will not only affect the use of the target subject, but also shorten life of the service, and even lead to catastrophic events. Therefore, NDT plays a crucial role in welding defect detection. Through the defect detection, it can be verified whether the weld quality of the tested target has reached the welding standard, so that it can find major defects in timely, remediate in time, and improve weld quality. In this process, the accuracy of weld inspection is particularly important.

In the traditional NDT process, the inspectors identify the defects of the weld by observing the X-ray scanning welding image, however, it is not easy to maintain the influence of personal emotions and fatigue, which make work efficiency very low and error-prone. In recent years, with the advancement of science and technology, the field of defect detection has been greatly developed, and some methods for recognizing defects have been proposed.

Welding defect detection of X-ray images is an important research subject in NDT. The study of welding defects classification is based on the use of the traditional methods on private dataset. The traditional method is to use mathematical morphology techniques to segment the image to extract the defect features<sup>[1]</sup>. Kasban, et al.<sup>[2]</sup> proposed a feature extraction method

that provides the basis for recognition. Domingo Mery, et al.<sup>[3]</sup> detected welding defects based on texture features. A method for the detection and classification using both texture features and geometrical features was presented by Valavanis, et al.<sup>[4]</sup>. Wang Gang, et al.<sup>[5]</sup> proposed a procedure for the automatic classification of welding defects, and the procedure includes three processes; image processing, feature extraction, and pattern recognition. The BSM method is used to extract defect area. And the fuzzy  $K$  nearest neighbor algorithm and the multilayer perceptron neural networks were used to classify separately. The artificial neural network (ANN) was used for classification tasks after feature extraction<sup>[6-7]</sup>. A different method to build the classifier based on image histogram and neural network was described by Abouelatta, et al. in Ref.<sup>[8]</sup>. The number of features was reduced by the principal components analysis (PCA) combined with ANN<sup>[9]</sup>. The support vector machine (SVM) classifier to classify based on texture features and morphological features were used for more efficient calculations<sup>[10]</sup>. The accuracy was improved based on the combination of PCA and SVM<sup>[11]</sup>. Baniukiewicz proposed a classifier consisting of neural network and fuzzy logic process<sup>[12]</sup>. Chen Benzhi, et al.<sup>[13]</sup> proposed to detect defects via sparsity reconstruction for weld radiographs. The welding area was located by detecting the regional characteristics of bands and the defects were divided into the categories of non-crack and crack roughly by analyzing gray column waveforms. According to the characteristics of the data, setting to manually design features is the current method.

Deep learning also provides a solution for welding defect detection. It will put a large number of welding images into the neural network, relying on deep learning to recognize defects and defect types, and gives the neural network people's experience and knowledge. Based on the basic structure of neural network, extracting features and related learning, the deep learning model achieves the effect of recognition. In the past researches in the field of NDT, some of recognition methods are only used to detect whether there are defects in the welding area<sup>[14]</sup>, and some are to test what kind of

defect types in the images. However, most of them could not to multi-type classify and localize effectively. So that, how to locate the defects in the X-ray images while classifying on WDXI is researched here.

Welding defects detection relying on deep learning is studied in this paper. Based on the Faster R-CNN model, a classifier is trained for welding defect classification and localization to solve a multi-classification problem. The established WDXI and the data preprocessing are introduced in the second part. The introduction and training of the model and the comparison of the training results are presented in the third part.

## 2 Dataset processing

All images are from the scene, where a database of 45 198 records is established. 16 950 images are selected of the database, where 13 638 records with images have complete and clear annotations and can be used for training and labeling. All the welding defects and defect assessments are conducted in strict accordance with the national standards of the People's Republic of China<sup>[15]</sup>. The dataset is divided into 7 types of defects Fig. 1, including round defects, bar defects, gas pores, lack of fusion, lack of penetration, slag inclusion, cracks, and no defects type, and the major welding defects are covered.

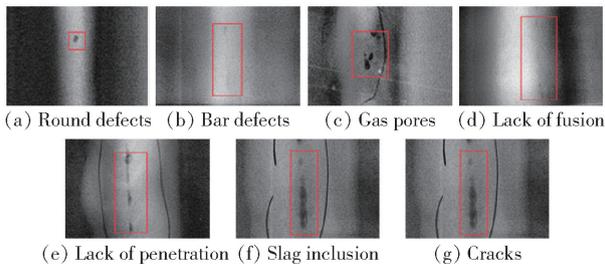


Fig. 1 7 types of defects

The original welding images are saved in the form of the X-ray films. The films are converted to digital images with a professional film scanner in Fig. 2. The WDXI can be divided into 7 types of defects Table 1. A total of 13 768 images are marked with defect types, including 130 no-defected images. The proportion of welding defect images of different defect types in the WDXI is quite different. The round defect and bar de-

fect are the most numerous types. And the unbalanced proportion of each defect has a great influence on the training result. The proportion of the number of images in each type of defects needs to be adjusted to balance the number of each type.

Table 1 Data distribution statistics

Defect type	Quantity
No label	3 182
No defect	130
Slag inclusion	969
Cracks	916
Gas pores	300
Bar defect	1 759
Lack of penetration	284
Lack of fusion	450
Round defects	8 960

### 2.1 Data augmentation

The defects of X-ray images are classified and labeled according to the corresponding testing reports. According to the number of images for each defect type, 1 000 images are randomly selected for each type of round defects and bar defects. All the images are selected for other defects types. And a new dataset is got from WDXI. Further image preprocessing and welding defect location annotation are performed based on this dataset.

There are 4 919 images for training dataset, including 7-type defects. For the defect types less than 1 000 images, data augmentation methods are used for increasing the number of images, for examples, randomly flipping, randomly clipping, randomly rotating, so that the number of images can reach 1 000. In this way, 7 000 images of dataset from WDXI are get. In the dataset, 10% of images are divided into testing set. In addition, 10% of the training set are divided into validation set. 700 images of test set, 630 images of validation set and 5 670 images of training set are obtained. And training set, validation set, and test set do not overlap with each other.

### 2.2 Image pre-processing

#### 2.2.1 Extract welding area

Firstly, the blanks of images producing when the

films are converted into digital images are cut in Fig. 2. Then, the welding area is extracted. The welding area is only a part of the whole image, and there is some marking information on the one side of it, such as film numbers. The texture features of the welding area are obviously different in both sides. According to the texture features of the image, the effective welding area can be extracted by comparing the average gray value, the average contrast value per unit area and the measures of texture features. There is a good method to extract welding area based on texture features<sup>[16]</sup>. According to the characteristics of the image, an improved method is put forward as follows.

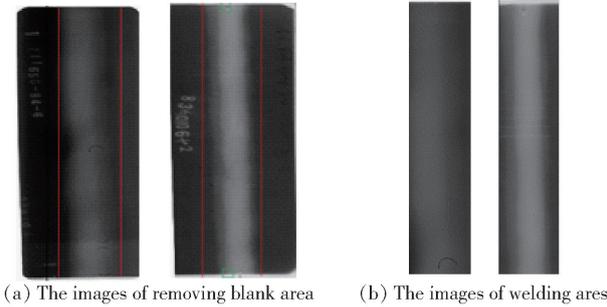


Fig. 2 The comparison of welding area before and after extraction

1) Set the width of the image is  $w$ , and the height is  $h$ . The average gray value  $m$  and the average contrast value per unit area  $c$  are calculated according to the following formulas.

$$m = \sum_{i=0}^{l-1} z_i P(z_i), \quad c = \frac{1}{wh} \sum_{i=0}^{l-1} (z_i - m)^2 P(z_i) \quad (1)$$

Where  $l$  is the range of possible grayscale values in the image,  $z_i$  is the pixel value within the range of  $l$ .  $P(z_i)$  represents the probability value of the  $z_i$  in the form of frequency of the image.

2) A sliding window of size  $(w_1, h_1)$  is used for calculating the local average gray value  $m_1$  and the local average contrast value per unit area  $c_1$ , where the  $w_1$  is 50 and the  $h_1$  is 100. The sliding window slides from the center of the top of the image in Fig. 2, and it moves to the left and right with 50 spacing, respectively. The sliding window will stop slide when the following conditions are met:  $m_1 < m$  and  $c_1 > 2c$ . The left boundary of the sliding window is the left boundary of the current weld area at the final position of the sliding

to the left, named  $b_{l,i}$ . And the right boundary of the sliding window is the right boundary of the current weld area at the final position of the sliding to the right, named  $b_{r,i}$ .

3) The sliding window is move down the distance of  $h_1$  and repeat the second step. And it will not stop until the sliding window is at the bottom of the image. There will be getting multiple left boundary values and right boundary values. The minimum left boundary value is taken as the left boundary value of the weld area  $b_l$ , and the maximum boundary value is taken as the right boundary value of the welding area  $b_r$ .

4) The boundary value is extended to ensure a complete welding area:  $b_l = b_l + 0.1b_l$ ,  $b_r = b_r + 0.1 \times (w - b_r)$ . The welding area with the size of  $[b_l, b_r, 0; h]$  is cropped. The lines obtained by the improved method are the boundaries of welding area in Fig. 2.

### 2.2.2 Image enhancement process

When all images are cut well, the result of inspection shows that the welding area portion of each image of dataset has been successfully extracted. And the AHE is used to enhance image and the DMB, the Gaussian blur (GB) and the digital subtraction technology (DST) are used separately to reduce image noise. The AHE improves effectively brightness and contrast, enhances image quality, and enhances the detail display of welding defect features. As shown in Fig. 3, it is obvious that the image processed by AHE is fresher, and the features are more obvious by contrasting the images of before and after processing.

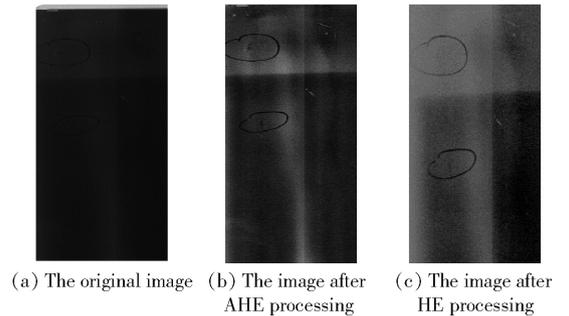


Fig. 3 Tailored from original image

The histogram equalization (HE) uses the same transformation derived from the image histogram to transform all pixels and the method works well when

the distribution of pixel values is similar throughout the image. However, the welding area is a bright strip-shaped area in the image, and it is significantly brighter than the areas on both sides. The HE cannot achieve good enhancement effect in the image. Since the AHE is based on local pixel region, it is used for image enhancement here.

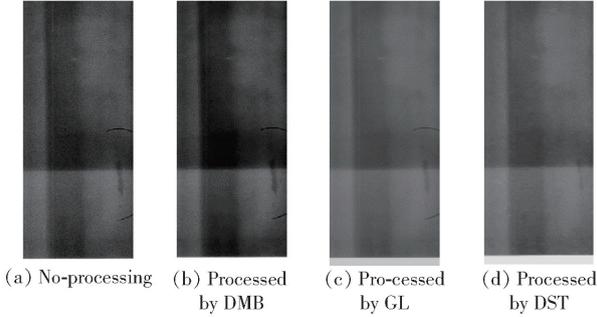


Fig. 4 A sample of image noise reduction processing

The traditional method, mathematical morphology, is used to simulate the background of welding area by selecting appropriate structural elements. The DST is used to remove the influence of background complexity, reduce the interference of the background noise and increase the difference between background and the welding defects. And the subtracted images are not of binaryzation to avoid excessive information loss. However, the shape and size of structural elements are mainly selected by experience at present, which are not universal and adaptable. Hence, the quality of some images is reduced after the processing.

The comparison of three methods for image noise reduction processing is shown as Fig. 4. Median blur is a kind of nonlinear filtering, which is excellent in processing impulse noise, and can effectively protect the edge information of images. It is obvious that the DMB is more effective than the GB. The DST reduces the impact of the background to some extent, but it also produces new noise in some X-ray images of WDXI. The effects of different noise reduction methods will be reflected in next experiments results.

### 3 Experiment

The purpose of the experiment is to train the classifier which can identify and position the welding de-

fects in the image. The potential welding defect types of the input welding image are localized in the image and classified. The faster R-CNN model<sup>[17]</sup> is composed of two parts, and the first part is a deep full convolution neural network for feature extraction and proposal region, the second part is Fast R-CNN<sup>[18]</sup>. Although the model is not the fastest, its mean average precision (mAP) is higher than single shot multiBox detector (SSD)<sup>[19]</sup>. When inputting the pre-processed welding image, and the model predicts the welding defect type and ground bounding box of the defects.

Some small amount of noise is enlarged by the AHE while improving the image contrast and more details. Therefore, in the process of image preprocessing, after AHE processing, the GB or the DMB is used to reduce noise.

#### 3.1 Training in Faster R-CNN

It is insensitive to the size of the input image in Faster R-CNN model. So that, there is no need to tailor the welding images. In Faster R-CNN, the pre-trained CNN model is used as the feature extractor. The ResNet<sup>[20]</sup> is chose to extract the image feature map. Since the image is grayscale image, the input layer of ResNet model is changed to operate on single channel grayscale image. The modified ResNet is used for multi-classification. Then, the trained ResNet will be the part of Faster R-CNN model.

There are some main parameter settings of the model. The initial learning rate is 0.000 1, learning rate decay factor is 0.001 and the learning rate decay exponentially. The Momentum algorithm is selected to optimize the loss function with  $\beta = 0.95$ . For the other parameter settings, for examples, initialization of weights by a zero-mean Gaussian distribution with standard deviation 0.01 or 0.001, the IoU threshold setting with 0.7 and 0.3.

The loss function is composed of RPN loss function and RCNN loss function, which consists of RPN regression loss function, RCNN regression loss function, RPN classification loss function, RCNN classification loss function, and extra robust loss function smooth L1 for regression function. There are the loss function curves in training Fig. 5. The loss function is

defined as:

$$L(\{p_i, t_i\}) = \frac{w_1}{N_{\text{rpn\_cls}}} \sum_i L_{\text{rpn\_cls}}(p_i, p_i^*) + \frac{w_2}{N_{\text{rpn\_cls}}} \sum_i L_{\text{rpn\_cls}}(p_i, p_i^*) + \frac{\lambda w_3}{N_{\text{rpn\_reg}}} \sum_i p_i^* L_{\text{rpn\_reg}}(t_i, t_i^*) +$$

$$\frac{\lambda w_4}{N_{\text{rpn\_reg}}} \sum_i p_i^* L_{\text{rpn\_reg}}(t_i, t_i^*) \quad (2)$$

In this formula, the weights,  $w_1$ ,  $w_2$ ,  $w_3$ , and  $w_4$ , of each loss function are 1. It is the same as the loss function in the original paper. The separation is to better control the optimization process of loss function.

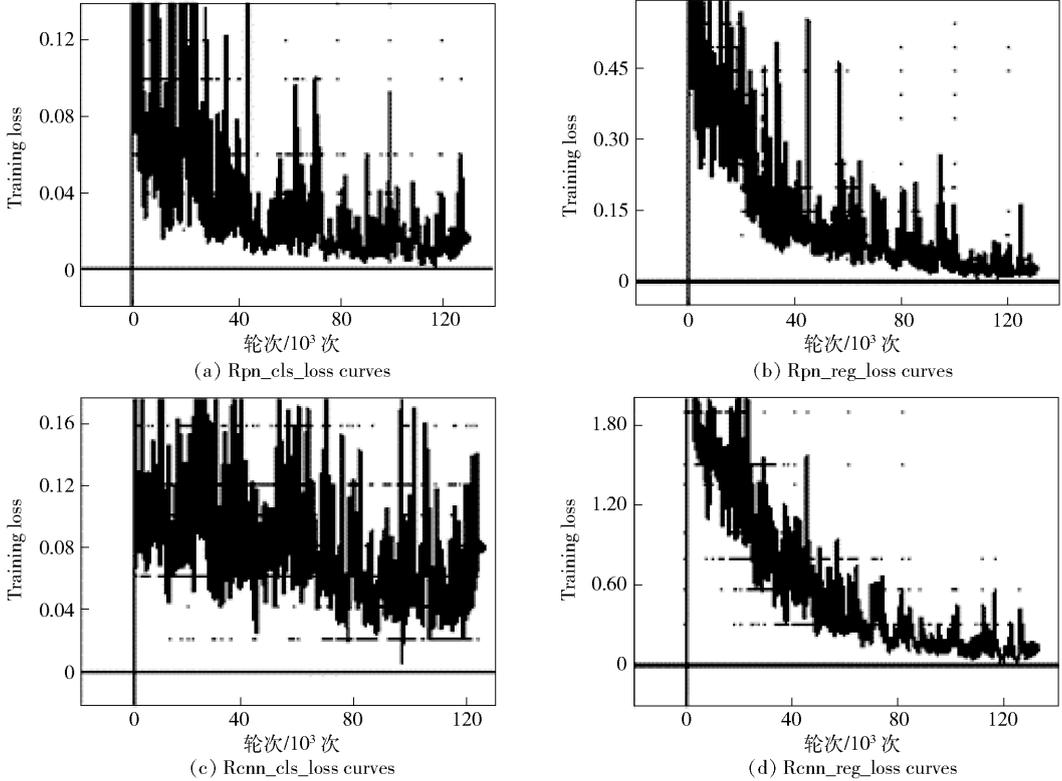


Fig. 5 The loss function curves

An example of the training process is shown in Fig. 6. The left side is the defect ground truth bounding box in the image annotation information of welding defect area. The right side is the proposal boxes with scores proposed which have been filtered by non-maximum suppression (NMS). In the middle is the Fast R-CNN's prediction of the ground bounding box by calculating these proposal boxes.

Using the same data set and partition, there are some experiments to show that the WDXI can be used in Faster R-CNN model and it is feasible for welding-defect detection based on Faster R-CNN model. The results with corresponding pre-processing are listed in Table 2.

Obviously, when the preprocessing methods are AHE and DMB, the mAP trained in faster R-CNN

Table 2 Detection result

Method of pre-processing	Model	mAP/%
no-pre-processing	4-conv networks	28.0
WWWL + GB	4-conv networks	46.6
no-pre-processing	Faster R-CNN	31.2
HE + GB	Faster R-CNN	26.2
AHE + DST + DMB	Faster R-CNN	47.4
AHE + DMB	Faster R-CNN	58.6
AHE + GB	Faster R-CNN	46.3
AHE + GB	SSD	38.2

model reaches 58.6%. According to comparing the score of mAP, it confirms some conclusions. In image preprocessing, the AHE is more effective than HE for image enhancement, the DMB is more excellent in processing impulse noise and protecting the edge informa-

tion of image. And there is no overall enhanced the quality of features after adding the DST processing.

In this experiment, the X-ray welding images of WDXI have the characteristics of high noise, low contrast, complex background area and small proportion of defect area in welding area. The characteristics make extra challenges for detection. The results of welding defect detection for X-ray images in trained Faster R-CNN model are showed in Fig. 6. In the process of object detection, the model recognizes the defect area and classifies subsequently defect types in it with the probability. The bounding boxes are obtained from the results predicted by the trained model. And the red bounding boxes are the actual label boxes for the welding defect area in images.

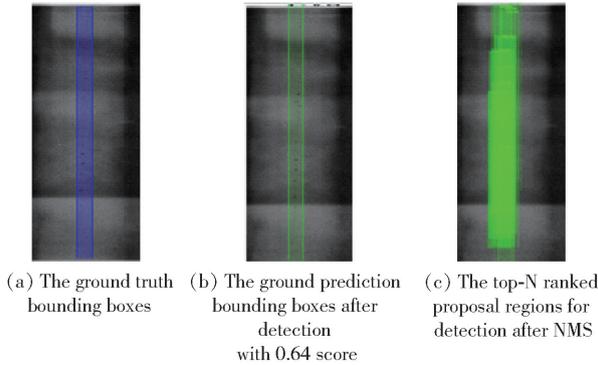


Fig. 6 The bounding boxes on training

In Table 4, there is the detailed information for the prediction results and the corresponding annotation information for the images in Fig. 7. The bbox represents the predicting ground bounding box, the gbox represents the ground bounding truth box, the prob represents the predicted probability value, prob\_label represents the predicted welding defect type and the true\_label represents the actual welding defect type. It can be seen that locating the welding defect area is accurate in trained model. In the negative example, the model gives the wrong judgment to the type of the welding defect in welding defect area. When the defect features in the image are slight, the model may make some mistakes in the classification of the welding defects. In the circumstance, it is easy to determine the defect type as other types with similar defect shape. This is a problem that needs to be optimized in the model optimization

work in the future, for improving the performance of the model.

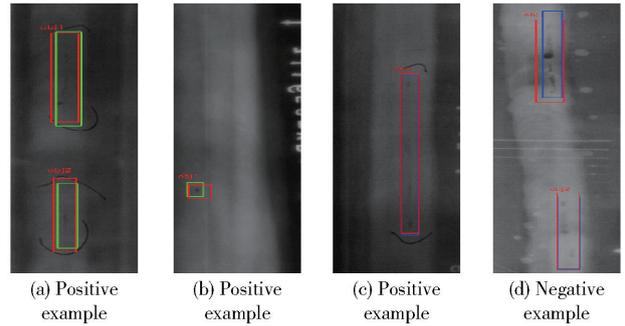


Fig. 7 Examples of object detection results on Tailored X-ray images test set

The testing results show that the performance of DMB processing with a sliding window sized  $7 \times 7$  is better than GB's in terms of noise reduction. And Faster R-CNN model is better than SSD model with the same preprocessing on WDXI. The best experiment result is showed in Table 3. The model has a better effect on detection of bar defect and cracks, which have large defect feature. Even though the data augmentation is performed for the not balanced different defect types, the imbalance problem inevitably affects the performance of the trained model. In general, preprocessed by AHE and DWB, the result is well trained on Faster R-CNN, and the mAP is 58.6%, which is acceptable in the exploration stage.

Table 3 Theap of the best experiment result

The type of defects	Average precision/%	mAP/%
bar defect	75.3	-
gas pores	53.2	-
round defect	60.0	-
lack of penetration	47.5	-
slag inclusion	60.8	-
lack of fusion	47.3	-
cracks	66.1	-
-	-	58.6

## 4 Conclusion

The deep learning algorithm Faster R-CNN is used for welding defect detection on WDXI, which is different from traditional methods and this experiment makes

**Table 4 The predicting information of the images in Fig. 7**

Obj	Bbox	Gbox	Prob	Prob_label	True_label
(a) obj1	[125, 386, 193, 933]	[111,393, 187, 912]	0.992 9	cracks	cracks
(a) obj2	[128,1 269, 183,164]	[118,1 236, 179,1 666]	0.980 9	crakes	cracks
(b) obj1	[45, 1 275, 97, 1 359]	[50, 1 290, 125, 1 368]	0.679 5	Round defects	Round defects
(c) obj1	[269,659, 336, 1 597]	[268,645, 337, 1 588]	0.819 2	lack of penetration	lack of penetration
(d) obj1	[194, 1 655, 271, 2 375]	[167, 1 728, 276, 2 417]	0.946 7	lack of penetration	slag inclusion
(d) obj2	[252, 3 202, 338, 3 810]	[251, 3 168, 337, 3 801]	0.797 4	gas pores	slag inclusion

an exploration in specific defect detection domain. By means of the deep learning model, the location of the welding defect area and the potential welding defect types with the probability in it are given to the inspector to help inspect the X-ray images. The image enhancement and noise reduction are vital to improve the performance of the training. The experiments show that the methods of AHE and DMB achieve a good effect on WDXI. The method, according to the average gray value and the average contrast value per unit area of the texture features, can extract the welding defect area well.

Experiments indicate that the WDXI has excellent research value in the field of NDT. The improved method of tailoring the image can successfully extract the welding area. The WDXI is available for academic research.

This paper aims at proposing an object detection system of welding defects based on faster R-CNN working on solving the problem of multi-classification classification and localization. The trained classifier reaches the 58.6% of mAP. It is acceptable in the experimental exploration to try different ways from traditional methods.

How to improve the performance of model detection and the application in actual work of the inspectors is the work for future work. Firstly, the work should solve the imbalance of different defect samples and the accuracy of labeling defect types and location. Then, more methods are tried to reduce the interference of complex background to defects, and improve the quality of feature map. Finally, optimize the Faster R-CNN and make the experiments with other neural networks and relevant algorithms according to characteristics of

the images.

### 参考文献:

- [1] Boaretto N, Centeno T M. Automated detection of welding defects in pipelines from radiographic images DWDI [J]. NDT & E International, 2017, 86: 7-13.
- [2] Kasban H, Zahran O, Arafa H, et al. Welding defect detection from radiography images with a cepstral approach [J]. NDT & E International, 2011, 44(2): 226-231.
- [3] Mery D, Berti M A. Automatic detection of welding defects using texture features [J]. Insight-Non-Destructive Testing and Condition Monitoring, 2003, 45(10): 676-681.
- [4] Valavanis I, Kosmopoulos D. Multiclass defect detection and classification in weld radiographic images using geometric and texture features [J]. Expert Systems With Applications, 2010, 37(12): 7606-7614.
- [5] Wang Gang, Liao T. Automatic identification of different types of welding defects in radiographic images [J]. NDT & E International, 2002, 35(8): 519-528.
- [6] Kumar J, Anand R S, Srivastava S P. Flaws classification using ANN for radiographic weld images [C] // 2014 International Conference on Signal Processing and Integrated Networks (SPIN). Amity: IEEE Press, 2014: 145-150.
- [7] Zapata J, Vilar R, Ruiz R. Automatic inspection system of welding radiographic images based on ANN under a regularisation process [J]. Journal of Nondestructive Evaluation, 2012, 31(1): 34-45.
- [8] Abouelatta, Ossama, et al. Classification of welding defects using gray level histogram techniques via neural network [J]. Mansoura Engineering Journal (MEJ), 2014, 39: M1-M13.
- [9] Vilar R, Zapata J, Ruiz R. An automatic system of classification of weld defects in radiographic images [J]. NDT & E International, 2009, 42(5): 467-476.
- [10] Wang Xin. Recognition of welding defects in radiograph-

- ic images by using support vector machine classifier [J]. *Research Journal of Applied Sciences, Engineering and Technology*, 2010, 2(3): 295-301.
- [11] Mu Weilei, Gao Jianmin, Jiang Hongquan, et al. Automatic classification approach to weld defects based on PCA and SVM [J]. *Insight - Non-Destructive Testing and Condition Monitoring*, 2013, 55(10): 535-539.
- [12] Baniukiewicz P. Automated defect recognition and identification in digital radiography [J]. *Journal of Nondestructive Evaluation*, 2014, 33(3): 327-334.
- [13] Chen Benzhi, Fang Zhihong, Xia Yong, et al. Accurate defect detection via sparsity reconstruction for weld radiographs [J]. *NDT & E International*, 2018, 94: 62-69.
- [14] 王明泉, 杨静, 李志刚, 等. 薄壁焊缝 X 射线图像缺陷的自动提取与分割 [J]. *计算机工程与应用*, 2007, 43(33): 237-239, 245.  
Wang Mingquan, Yang Jing, Li Zhigang, et al. Automatic defect extraction and segmentation in X-ray images of welding seam [J]. *Computer Engineering and Applications*, 2007, 43(33): 237-239, 245.
- [15] GB/T 3323—2005, Radiographic examination of fusion welded butt joints in steel [S]. BSI British Standards. DOI: 10.3403/30307847u.
- [16] 张晓光, 孙正, 胡晓磊, 等. 射线检测图像中焊缝和缺陷的提取方法 [J]. *焊接学报*, 2011: 0253-360X.  
Zhang Xiaoguang, Sun Zheng, Hu Xiaolei, et al. Extraction method of welding seam and defect in ray testing image [J]. *Transactions of the China Welding Institution*, 2011: 0253-360X.
- [17] Ren Shaoqing, He Kaiming, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [18] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE Press, 2015: 100-105.
- [19] Liu Wei, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector [M] // *Computer Vision-ECCV 2016*. Cham: Springer International Publishing, 2016: 21-37. DOI: 10.1007/978-3-319-46448-0\_2.
- [20] He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Deepresidual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE Press, 2016: 105-110.