Paper Defects Recognition Based on Deformable Convolution

Yun-hui Qu[1,2][†], Wei Tang[3], Bo Feng[4]

ABSTRACT

There are some problems in traditional paper defects classification, such as the poor gen-eralization performance, less types of recognition, and insufficient recognition accuracy. The deep learning method provides a new scheme for paper defects classification. However, convolutional neural network has strict requirements for the size of the input image. This requires that in the process of practical engineering application, for the collected paper defect images to be classified, the area containing paper defect must be segmented during preprocessing, and then the size of the paper defect area must be adjusted to meet the input requirements of the adopted classifier. To solve the above problems, the two-stage target detection network Faster R-CNN (Region-Convolutional Neural Network) was used in paper defects recognition to solve the problem of the size requirements of the input image; In addition, the deformable convolution layer was added after the traditional convolution layer to learn the characteristics of paper defects more efficiently and accurately, so as to improve the accuracy and accuracy of paper defects recognition and classification; Finally, the deformable RoI (Region-of-Interest) pooling layer was used to replace the RoI pooling layer of classic Faster R-CNN to locate and classify the paper defects area more accurately. Experiments show that the proposed algorithm has a further improvement in accuracy and scalability compared with the previous algorithm.

---

1 Computer Teaching and Research Section, Xi'an Medical University, People's Republic of China, Professor
2 Department of Electric and Control Engineering, Shaanxi University of Science & Technology, People's Republic of China, Student
3 Department of Electric and Control Engineering, Shaanxi University of Science & Technology, People's Republic of China, Professor
4 Department of Electric and Control Engineering, Shaanxi University of Science & Technology, People's Republic of China, Student
† Corresponding Author: E-mail: nan-nan_1951@163.com (Address: Computer Teaching and Research Section, Xi'an Medical University, Xi'an, Shaanxi, 710021, People's Republic of China)

1. Introduction

In the process of paper production, the paper defect is the surface defects such as dirty spots, holes, folds, scratches, dust and cracks. The appearance of paper defects will have a negative impact on the subsequent use, especially on aerospace paper, electrolytic capacitor paper, wallpaper base paper and other special paper with high added value in the later period, which will bring huge economic losses, so it is necessary to carry out accurate diagnosis and timely treatment. The web diagnosis technology is to judge whether the paper contains paper defects by collecting the paper image online with industrial camera. If the paper images containing defects, the further classification and recognition are carried out. Generally speaking, the web diagnosis process can be divided into three main stages: paper image acquisition and preprocessing, paper defects online detection and paper defects recognition or classification.

Paper defects recognition is a key step in the whole paper defect diagnosis process. Accurate recognition of various paper defects is of great significance for finding out the causes of paper defects and early warning of paper machine malfunction.

In the paper defects recognition, the most research was to use artificial neural network and convolutional neural network to establish classifier. Ni et al.[1] used BP (Back Propagation) neural network to establish a paper defects classifier to distinguish four kinds of paper defects: holes, dirty, folds and cracks. Wei et al.[2] improved the problem of slow convergence speed of BP neural network, and used the total error generated by learning samples to adjust the weight and improve the speed of BP algorithm. Support vector machine (SVM), two-dimensional wavelet transform, classification theory, radial basis function neural network (RBFNN) and other applications in paper

defects classification have been studied.[3–5] These algorithms needed to extract the paper defects features first, take the extracted paper defects features as the input of the classifier, and output the classified results. Therefore, in order to obtain a good classification effect, a large number of different types of paper defects features should be selected and designed a powerful multi paper defects classifier. Therefore, for each paper defect that needs to be identified and classified, we need to research and extract its features and transform the classifier, so it is difficult to improve the identification type and efficiency.

In recent years, with the research and popularization of deep learning and convolutional neural network, many scholars had proposed paper defects classifiers based on convolutional neural network. But convolutional neural network had strict requirements for the size of input image, which required that in the process of practical engineering application, the area containing paper defects must be segmented during preprocessing, and the size of the paper defect area must be adjusted to meet the input requirements of the adopted classifier. For example, VGG16 convolutional neural network requires the input image size to be 224×224 pixels, while the size of a frame image collected by the industrial camera is 4096×1048 pixels, about 84 times the input image size required by VGG16. In the process of normalization, there would inevitably be problems such as information loss or object deformation. Once there was an error in the paper defects pretreatment process, it would inevitably lead to the error of paper defects classification in the later stage.

In view of the above problems, based on the classical object detection network Faster R-CNN (Region-Convolutional Neural Network), a paper defects recognition method was proposed based on deformable neural network. Firstly, the two-stage object detection network Faster R-CNN was used

as the basic network to solve the problem of the size requirements of the input image. Secondly, the deformable convolution layer was added after the traditional convolution layer of VGG16 for feature extraction, so as to improve the accuracy of paper defects recognition and classification. Finally, the deformable RoI (Region-of-Interest) pooling was used to replace the RoI pooling of the classic Faster R-CNN to locate and classify the paper defects area more accurately.

2. Detection Network Faster R-CNN

At present, the paper defect diagnosis system of paper production line is mostly divided into two independent stages: paper defects detection on-line and paper defects classification. In the paper defects detection on-line stage, it only detects whether the paper image contains paper defects, and does not distinguish the types of paper defects. In the paper defects classification stage, the types of paper defects contained in the paper are classified. At present, the classification process is mostly completed by convolutional neural network. Due to the limitation of CNN model, the paper images need to be preprocessed and normalized. At the same time, each paper image can only contain one paper defects, which greatly limits its generalization ability. Therefore, the two-stage image object detection network Fast R-CNN was used as the basic network, and proposed a paper defects recognition algorithm with stronger generalization ability, which could classify paper defects and identify paper defects areas at the same time.

2.1 Object detection

Object detection is a computer vision technology which uses algorithms to search for objects of interest in images.[6] Traditional object detection included preprocessing, window sliding, feature extraction, feature selection, classification, post-processing and so on.[7] The traditional feature extraction had poor generalization ability and low accuracy. Because the convolutional neural network (CNN) has the ability of feature extraction, selection and classification, so CNN can directly be used to complete feature extraction, classification and regression according to the extracted features in the process of object detection.

The object detection algorithm based on deep learning can be divided into single-stage object detection algorithm and two-stage object detection algorithm. The two-stage detection algorithm treats object detection as a classification problem, that is, first generate object candidate regions, and then classify and calibrate the candidate regions to obtain the final detection results.[8] Among the two-stage object detection algorithms, the region based convolutional neural network algorithms were the most widely used object detection algorithm at present, such as R-CNN, Fast R-CNN, Faster R-CNN, etc. R-CNN series algorithms combined region proposal with convolutional neural network (CNN), and used convolutional neural network to classify candidate regions and judge them as background or objects.[9] Compared with the six steps of the traditional object detection algorithms, the region based convolution neural network object detection algorithms only had three steps: generating candidate regions, classifying candidate regions and post-processing, which had strong generalization ability and high object detection accuracy.

2.2 Faster R-CNN

Fast R-CNN is an improved object detection network structure based on R-CNN and Fast R-CNN, aiming at the problems of large disk space occupation, waste of resources caused by repeated feature extraction, slow training and testing speed.[10-12]

Fast R-CNN object detection network was mainly composed of region proposal networks (RPN) for generating candidate regions and Fast R-CNN for classification and boundary regression.[13-14] The two parts share the same convolutional neural network to extract image features. In this way, the detection time of object candidate region is greatly shortened, the speed of object detection is improved, and which is more suitable for the real-time diagnosis process of industrial generation line.

The network structure of Fast R-CNN was shown in Fig. 1.

The realization of Faster R-CNN function was mainly completed by the following networks:

1) Feature extraction network

The feature extraction part of Faster R-CNN was used to extract the feature map of the image. Its structure was the same as that of CNN network, including a series of convolution and pooling operations. Therefore, the feature extraction of this part can be directly completed by using the classical network model. In the algorithm proposed in this paper, the feature extraction of this part is improved based on VGG16.

2) Region proposal network (RPN)

RPN was used to produce regional candidate boxes. It determines the object and background through softmax, and continues to identify and classify the regional candidate boxes determined as the background.

The input of RPN can be an image of any size. The output is a batch of rectangular region proposals, and each region corresponds to a object score and location information.

3) RoI poolling

RoI pooling synthesized the feature map generated by the first two parts and the information of the candidate box. The coordinate position of the candidate box in the input image was mapped to the last obtained feature map, the corresponding position in the feature map is pooled, and the output was connected as an input to the last classification layer for classification.
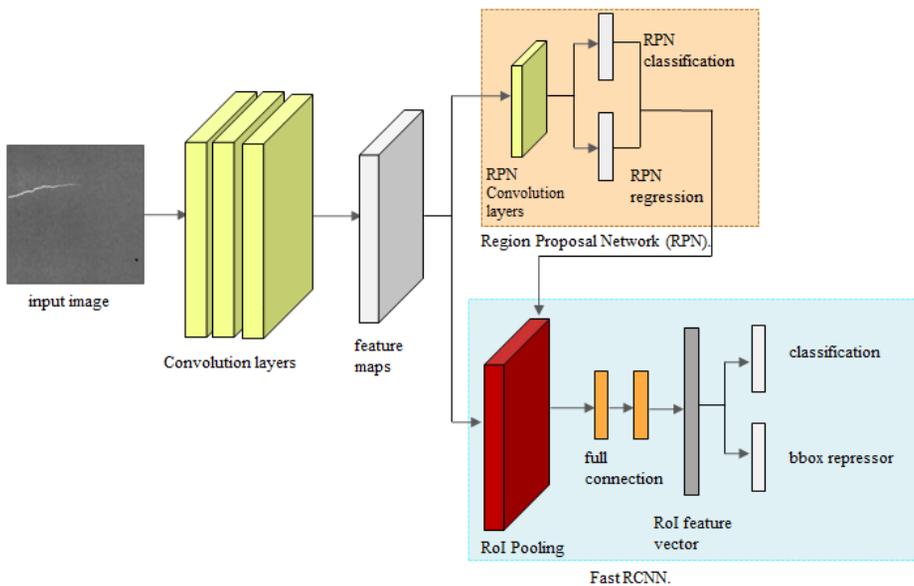

Fig. 1. Network structure of Faster R-CNN.

4) Classification layer

Used to determine the category of candidate boxes. At the same time, the classification layer can be connected to calibrate the accurate position of the candidate frame.

3. Paper Defects Recognition based on Deformable Convolution Neural Network

3.1 Deformable convolution

In the traditional convolution neural network model, the size of convolution kernel is mostly fixed 5×5, 3×3 or 1×1. For complex or small image features, this kind of convolution kernel may lose key feature information. As described in reference,[15] the convolution kernel size of VGG16 is 3×3. Therefore, if the paper defect area was small dirty spots, holes or slender folds and cracks, the convolution kernel would not be able to perceive and extract its features, and resulting in misclassification. Based on this, Dai[16] proposed the concept of deformable convolution in 2017, which was used to improve the fixed geometry of convolution neural network model and improve the spatial information modeling ability of traditional convolution neural network,[17] so as to solve the problems such as the limited size of convolution kernel of traditional convolution neural network.

An offset was added to the corresponding position of each sampling point in the convolution kernel. Through these offsets, the convolution kernel could sample randomly near the current position, and was no longer limited to the previous regular lattice points. The expanded convolution operation is called deformable convolution.

3.2 Realization of deformable convolution

The implementation of deformable convolution is to introduce two new modules into convolution neural network to enhance the modeling ability of original CNN for geometric transformation. The two modules are deformable convolution and deformable RoI pooling. Both are based on the idea of adding spatial sampling location in the module, which has additional offset and learns the offset of the target task without additional supervision. The new module can easily replace the ordinary peers in the existing CNN, and can easily carry out end-to-end training through standard back propagation to produce a deformable convolutional network.

3.2.1 Deformable convolution

Deformable convolution means that the convolution kernel adds an additional parameter direction parameter to each element, so that the convolution kernel can be extended to a large range in the training process. That is, the two-dimensional offset is added to the grid sampling position of the traditional convolutional neural network, so that the sampling grid can be deformed freely. The offset can be learned from the input feature map by adding an additional convolution layer, and its size depends on the input feature.

The structure of deformable convolution is shown in Fig. 2.

In Fig. 2, the above bypass path learns the size of offsets through a convolution layer (conv). In the figure, the offset field is an additional convolution layer, and the number of channels is twice the size of the convolution kernel. For example, for 3×3, if the size of the convolution kernel is 9, the number of channels is 18.

The mathematical expression of deformable convolution is: let R be a sampling matrix with p_0 as the center point, the mathematical expression of deformable convolution of any point y (p_0) on the input characteristic graph is shown in Eq. (1).

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_{ij}) \tag{1}$$
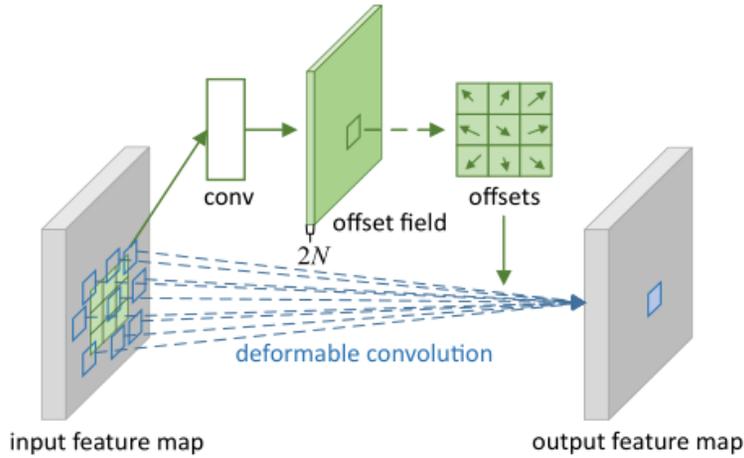
Fig. 2. Deformable convolution.[16]

Where,

$w(P_n)$: The weight corresponding to each sampling point.

ΔP_{ij}: Deformable convolution is a learnable offset added at each sampling point position of the standard convolution.

3.2.2 Deformable RoI pooling

Deformable RoI pooling is to add an offset to each sub area (bin) of the previous RoI pooling to move as a whole, so that it can adapt to the local positioning of objects with different shapes. Similarly, the offset can be learned from the input feature

map and RoI region. The structure of deformable RoI pooling is shown in Fig. 3.

The mathematical expression of deformable RoI pooling is shown in Eq. (2).

$$y(i,j) = \sum_{p \in bin(i,j)} x(p_0 + p + \Delta p_{ij}) / n_{ij} \qquad (2)$$

As shown in Fig. 3, the offset was obtained by bypassing the convolution layer. Firstly, the bypass performed RoI pooling on the input feature map to generate a new feature map. Then the full connection layer (FC) operation was performed on the new
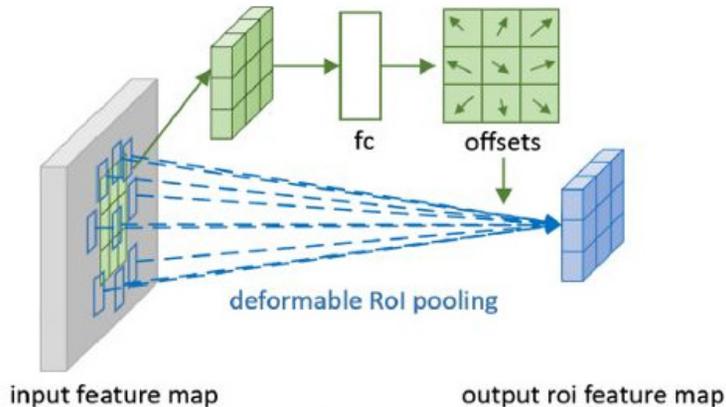

Fig. 3. Deformable RoI Pooling.[16]

feature map to output the normalized offset $\Delta \hat{p}_{ij}$. Δp_{ij} could be obtained by multiplying the normalized offset $\Delta \hat{p}_{ij}$ and the width and height (w, h) of RoI on the element wise product.

3.3 Paper defects recognition based on deformable convolution neural network

Aiming at the problem of low recognition accuracy of existing paper defects recognition algorithms, a paper defects recognition algorithm based on deformable convolutional neural network was proposed by combining the idea of deformable convolutional neural network with the classical object detection model Faster R-CNN.

Firstly, according to the characteristics of each paper defect image studied in the previous stage,[15] VGG16 convolution neural network model was used as the feature network, and a deformable convolution layer was added to its traditional convolution layer to better perceive the characteristics of different types of paper defects. Secondly, the region of interest (RoI) pooling layer in Faster R-CNN network was replaced by deformable RoI pooling to further improve the accuracy of detection area.[18]

3.3.1 Network structure design based on deformable convolutional neural network

1) Feature extraction network based on deformable convolution

In this algorithm, VGG16 was used as the basic network for paper defect feature extraction, in which the input image was the paper image in the paper defect image database. Secondly, two deformable convolution layers were added after the fourth and tenth layers (i.e. the second and fourth construction layers) of the 13 convolution layers of the original VGG16 convolution neural network. The third construction layer outputs the extracted texture features, and the fifth construction layer extracts the local characteristics of the object. Therefore, the deformable convolution was added before the two construction layers, and the geometric transformation with additional offset was used to effectively deal with the paper defect area, so as to make the extracted paper defect image features more accurate, and effectively improve the accuracy of paper defect classification.

The network structure of paper defect feature extraction based on deformable convolution is shown in Fig. 4.
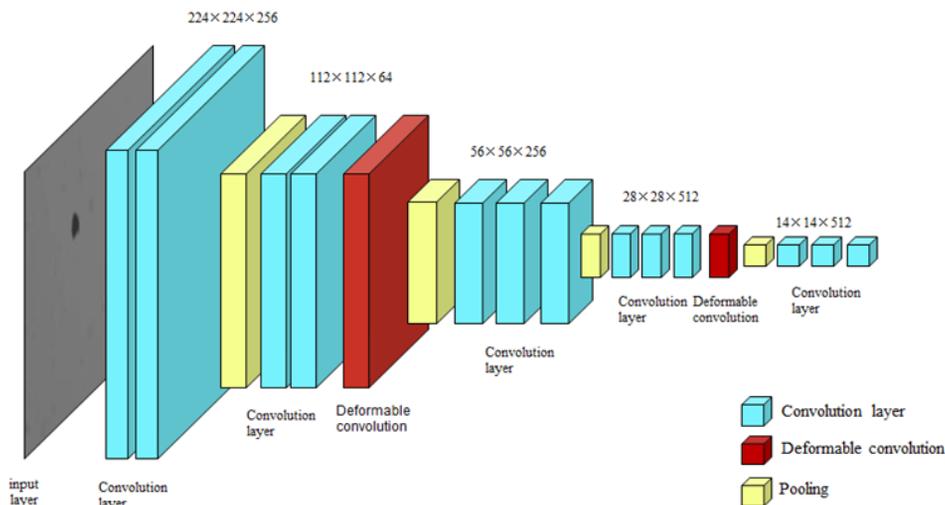

Fig. 4. Structure of deformable convolutional neural network.

2) Deformable RoI pooling

In this algorithm, the RoI pooling layer in the Faster R-CNN network (shown in Fig. 2) was replaced by the deformable RoI pooling. Because deformable RoI pooling has the same input and output forms as ordinary RoI pooling layer, the corresponding layers in the existing model can be directly replaced to more accurately locate the paper defect area in the image, so as to carry out more accurate classification.

3) Parameter setting of deformable structure

The structures of the added deformable convolution layer and the deformable RoI pooling layer were shown in Fig. 2 and Fig. 3 respectively. In Fig. 2, the additional convolution layer (offset field) for learning offset in the upper path and the FC layer in Fig. 3 were initialized with 0. The learning rate of the deformation layer was consistent with that of the layer.

3.3.2 Loss function

The deep convolution neural network learns the parameters through the back propagation of the error between the prediction results of sample data and the real mark. For classification tasks, commonly used loss functions include cross entropy loss function, high loss function, ramp loss function and center loss function. The ramp loss function and center loss function are generally used in the classification task with more sample noise. For the paper defect images, due to the simple acquisition environment and less noise, the cross entropy loss function was used in the research. For the classification task, its effect is generally better than the hinge loss function.

The cross entropy loss function is also called softmax loss function, which is described as follows:

There are N training samples in the classification task, which are divided into class C. The input feature of the i th sample of the classification layer is x_i, and its corresponding real mark is $y_i \in \{1,2,..., C\}$. The final output of the network (i.e. the prediction result of sample i) is $h=(h_1,h_2,...,h_c)^T$, the cross entropy loss function is shown in Eq. (3).

$$loss = -\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{e^{h_{y_i}}}{\sum_{j=1}^{C} -e^{h_j}}\right) \tag{3}$$

3.3.3 Network regularization

In order to prevent the "over fitting" phenomenon of the model and make the model have strong generalization ability, the "regularization" technology was used to control the complexity of the model, so that the model could not only perform well on the training samples, but also uses good classification results on new data and test sets.

Dropout is a network regularization method commonly used in the full connection layer of the current deep convolution neural network. It is an efficient integrated learning method for the depth model while constraining the network complexity. To a certain extent, it alleviates the complex cooperative adaptation between neurons and reduces the laziness between neurons, so as to avoid the phenomenon of over fitting. The principle of dropout was described as follows:

In the process of network training, the weight of each neuron in the whole connection layer is randomly reset to 0 with probability p (i.e. in the "inactivated" state). In the test phase, all neurons are activated, but the weight of each neuron needs to be multiplied by $1-p$, so that the weight in the training and test phases has the same expected value.

4. Results and Discussion

4.1 Experimental environment and parameter setting
4.1.1 Experimental environment

The experimental environment was Win10 operating system and Intel Core™ I7-7500u CPU, 8G Ram, 256G SSD. Matlab 2019b platform was used to build and train the model.

4.1.2 Parameter setting

1) Parameter initialization

In the process of network training, the initialization of network parameters determines the final performance of the network to a great extent. In order to get better training effect, this paper used the parameters of the pre-training model to replace the random initialization method. In the previous research work (as described in reference 15), the VGG16 model had been trained to converge by using the paper defect image. Therefore, in this work, the model training parameters were saved and loaded into the deformable convolution network for training to learn the deformation information of the convolution kernel.

2) Self supervised training selection optimizer

This paper used Adam self-supervised training to select the optimizer. Adam uses the first-order moment and second-order moment estimation of the gradient to dynamically adjust the learning rate of each parameter. Its advantage is that after bias correction, the learning rate of each iteration process will be limited to a certain range, so that the parameter update is relatively stable.

During the experiment, the basic learning rate of Adam optimizer was 0.001 and the dynamic variation range of momentum was 0.9 - 0.99.

3) Other parameter settings

Setting the maximum number of iterations of training to 5000; Batch size was 64. In the training, the learning rate was attenuated every 5000 steps, and the attenuation coefficient was 0.005.

4.2 Experimental results and analysis
4.2.1 Data acquisition

In this study, five kinds of paper defect images and no defect images were collected by the laboratory web inspection equipment.

Because the appearance of paper defects was a small probability event, the training and testing of convolutional neural network need a large number of data sets. Although the method of transfer learning and fine tuning depth convolution neural network can reduce the requirement of sample size, sufficient training data can avoid the problem of over-fitting. Therefore, in order to increase the scale of paper defect images data set, in the process of making paper defect image data set, mirror image, rotation and other operations (such as 90°, 180° and 270° rotation of the paper defect images, which can increase the paper defect images data set while maintaining the paper defect images characteristics) and multiple acquisition of the same paper defects under the interference of different light sources were used to expand the data set. Finally, 370 dirties, 340 holes, 280 bright spots, 350 folds, 360 cracks and 300 normal paper images were obtained as the paper defect images data set. The training set and test set were divided according to the ratio of 4:1.

4.2.2 Recognition effect and analysis of test samples

The trained model was tested with the test set data. Some samples and their recognition effects were shown in Fig.5. The number after the type represented the probability of belonging to this type. As shown in Fig. 5c, the identified area was crack 0.86, indicating that 86% of the probability of this area was crack. From the recognition
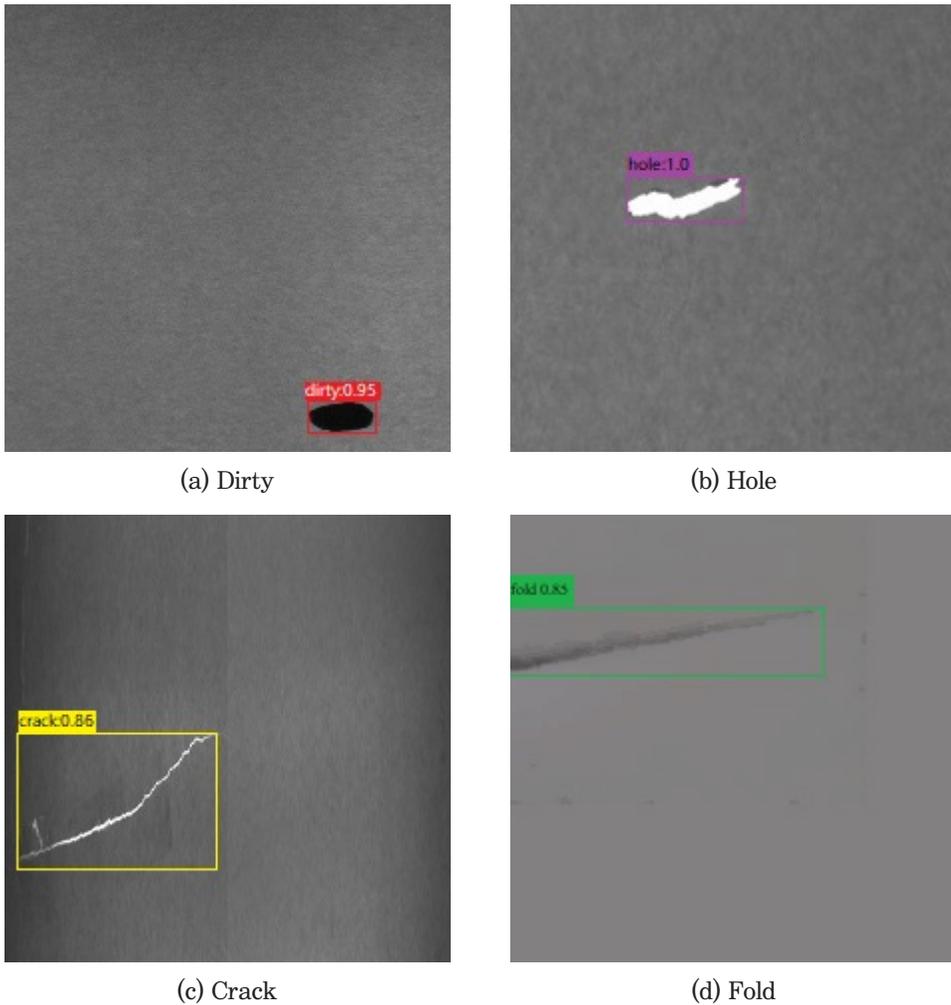
(a) Dirty

(b) Hole

(c) Crack

(d) Fold

Fig. 5. Various paper defect detection results.

results of four kinds of paper diseases and multi paper defects images given in Fig. 5, it could be seen that the classification correctness and posi-tioning accuracy of the algorithm proposed in this paper were relatively well.

Fig. 6 shows the recognition effect of multi paper defect image. As can be seen from Fig. 6, for multi paper defect images, the algorithm in this paper can accurately mark multiple paper defects and give accurate classification results.

Meanwhile, the paper defect images size shown in Fig. 5 was 224×224 pixels, but Fig. 6 was not 224

×224 pixels (the actual size of the two images shown in Fig. 6 was 689×516 pixels). It could be seen that the algorithm still had good recognition effect. So the algorithm proposed in this paper had strict requirements on the size of the input paper defect images.

During the experiment, the recognition of five kinds of paper defects and normal paper images without paper defects were statistically analyzed in the experiment. The results are shown in Table 1. The determination of normal paper image without paper defect was that after the region proposal
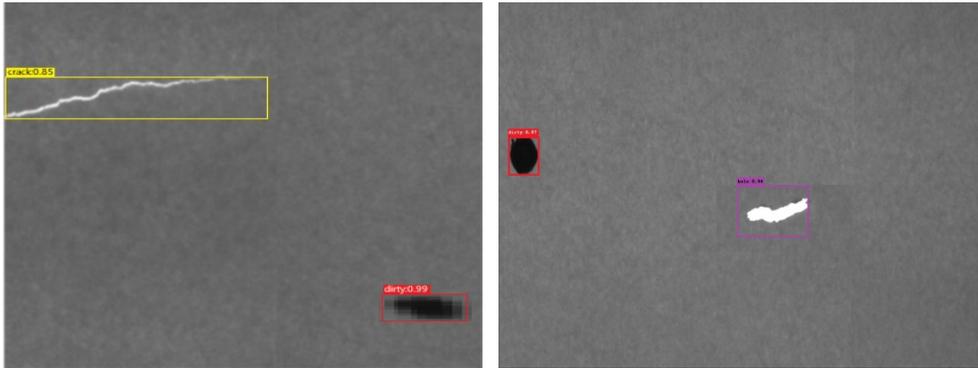
Fig. 6. Detection result of multiple paper defects.

Table 1. The classification results of the test samples

	Dirty spot	Hole	Bright spot	Fold	Crack	Normal paper	Correct number	Total	Accuracy (%)
Dirty spot	74	0	0	0	0	0	74	74	100
Hole	0	65	3	0	0	1	65	68	95.59
Bright spot	0	2	51	0	0	3	51	56	91.07
Folds	0	0	0	67	1	2	67	70	95.71
Crack	0	0	0	1	69	2	69	72	95.83
Normal paper	0	0	0	0	0	60	60	60	100
Total samples							386	400	96.50

network generates candidate frames, if all candi-date frames were determined as background, this image was considered as normal paper image without paper defect.

As can be seen from the data results in Table 1:

1) The algorithm proposed in this paper had a high recognition and classification accuracy, which could reach more than 90% for various types of paper defect.

2) For bright spot, due to its low contrast com-pared with paper background, the recognition accuracy of previous algorithms was low. But the classification accuracy of the algorithm proposed in this paper had increased to more than 91%, which was greatly improved com-pared with the traditional algorithm.

3) For holes, folds and cracks with obvious con-trast, the accuracy could reach more than 95%,

and the effect was very well.

In conclusion, the accuracy of the paper defects recognition algorithm based on deformable convo-lutional neural network proposed in this paper had been greatly improved compared with the original algorithm, and effectively solved the problem of low accuracy of the traditional paper defects rec-ognition algorithm.

4.2.3 Comparative analysis of simplified models

Since the main work of the proposed method was to change the existing Faster R-CNN structure so that it could be applied to the paper defects recog-nition, the effectiveness of the modified network structure was important to measure the work of this paper.

In order to further verify the effectiveness of the deformable convolution structure proposed in this

paper, the deformable convolution structure proposed in Fig. 4 was further improved for comparison. During the experiment, the first and second deformable convolution layers in the network shown in Fig. 4 were removed and the deformable RoI pooling layer was replaced by ordinary RoI pooling. The above simplified network model was trained in the same experimental method, and the classification accuracy and detection speed were tested on the test set. The comparison results were shown in Table 2.

It can be seen from the comparison of experimental results in Table 2:

1) Compared with the Faster R-CNN object detection network before adding deformable structure, although the detection speed was slow, the classification accuracy was improved by about 3 percentage points.

2) The deformable convolution neural network proposed in this paper had the highest accuracy in paper defects recognition and classification, and the detection speed was almost the same as that of removing a deformable convolution layer or replacing the pooling layer.

In the whole process of paper defect diagnosis, the recognition and classification of paper defect images were the last step, so the requirement for real-time was relatively low. At the same time, the occurrence of paper defect was a relatively small probability event. Therefore, after early rapid detection, the number of paper images sent to the last step for recognition and classification was relatively small. Based on the above two reasons, the recognition speed of the algorithm proposed in this paper was about 1 frame/second, that was, the time to detect an image was about 1 second, which basically meets the use of the actual production line.

5. Conclusions

Aiming at the low accuracy of paper defects recognition caused by the small image area or irregular shape of the paper defect image collected on the actual production line, a paper defects recognition method based on the two-stage image detection algorithm Faster R-CNN and deformable convolution was proposed. Due to the small area and irregular shape of the paper defects, the classification accuracy of Faster R-CNN in the paper defects recognition process was low, and the location was not accurate enough. Aiming at the above problems, two deformable convolutions were added after the traditional convolution layer to extract the characteristics of the paper defects more accurately. And then, the deformable RoI pooling was used instead of the common RoI pooling, which made the positioning more accurate. Experiments show that the proposed algorithm has a further improvement in accuracy and scalability compared with the previous algorithm.

Table 2 Results of ablation test

Model	Accuracy (%)	Recognition speed (Unit: frame / second)
Remove the first deformable convolution	95.45	1.2
Remove the second deformable convolution	95.47	1.7
RoI pooling	94.45	1.4
Faster R-CNN	93.75	3.3
Proposed	96.50	1.0

Acknowledgement

Literature Cited

1. Ni, J., Xu, J., and Hu, M.Y., Paper defects classifier design based on BP neural network, Transactions of China Pulp and Paper 25(2): 76-78 (2010).

2. Qu, Y. H., Tang, W., and Feng, B., Web inspection algorithm for low contrast paper defects based on artificial bee colony optimization, Journal of Korea TAPPI 52(2):43-51 (2020).

3. Qu, Y. H., Tang, W., and Wen, H., On-line detection and classification method based on background subtraction and SVM, Packaging Engineering 9(23):176-180 (2018).

4. Qu, Y. H., Tang, W., and Feng, B., Web inspection algorithm for low contrast paper defects based on artificial bee colony optimization, Journal of Korea TAPPI 52(2):43-51 (2020).

5. Li, G. M., Xue, D. H., and Jia, X. H., Paper defects classification based on multi-scale image enhancement combined with convolution neural network, China Pulp and Paper 8(37):47-54(2018).

6. Lu, Q. S., Research on object detection method based on deep learning, Beijing: Beijing University of Posts and Telecommunications 4:6 (2020).

7. Tian, H. L., Ding, S., and Yu, C. W., Research of video abstraction based on object detection and tracking, Computer Science 43(11):297-299 (2016).

8. Li, X. D., Ye, M., and Li, T., Review of object detection based on convolutional neural networks, Application Research of Computers 34(10):2881-2886, 2891 (2017).

9. Wu, X., Song, X. R., and Gao, S., Review of target detection algorithms based on deep learning, Transducer and Microsystem Technologies 40(02):4-7, 18 (2021).

10. Girshick, R., Donahue, J., and Darrell, T., Rich feature hierarchies for accurate object detection and semantic segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587 (2017).

11. Ren, S., He, K., and Girshick, R., Faster R-CNN: Towards real-time object detection with region proposal networks, 29th Annual Conference on Neural Information Processing Systems, pp. 91-99 (2015).

12. Ma, J. L., Chen, B., and Sun, X. F., General objects detection framework based on improved faster R-CNN, Journal of Computer Application 41(9):2712-2719 (2021).

13. Cai, Z. X., Li, R. X., and Dai, Y. D., Fabric defect recognition system based faster R-CNN, Journal of Computer Application 30(2):83-88 (2021).

14. Cheng, Y., Xia, L. Z., and Yan, B., A defect detection method based on faster RCNN for power equipment, Journal of Physics: Conference Series 1754(1):1884-2022 (2021).

15. Qu, Y. H., Tang, W., and Feng, B., Paper defects classification based on VGG16 and transfer learning, Journal of Korea TAPPI 53(2):5-14 (2021).

16. Dai, J., Qi, H., and Xiong, Y., Deformable convolutional networks, Proceedings of the 2017 IEEE International Conference on Computer Vision, pp. 764-773 (2017).

17. Zhu, X., Hu, H., and Lin, S., Deformable ConvNets v2: more deformable, better results,

32nd IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9300–9308 (2019).

18. Wu, S. M., Zhu, Y., and Wang, F., An electronic device container quality detection method based on cascade R–CNN, Computer and Modernization 2020(11):33–38, 46 (2020).