A Faster-RCNN Based Chemical Fiber Paper Tube Defect Detection Method

Yuzhou Shi
School of Aeronautics and Astronautics
Shanghai Jiao Tong University
Shanghai, China
Connor_syz@sjtu.edu.cn

Yuanxiang Li
School of Aeronautics and Astronautics
Shanghai Jiao Tong University
Shanghai, China
yuanxli@sjtu.edu.cn

Xian Wei
Quanzhou Institute of Equipment Manufacturing
Haixi Institutes, Chinese Academy of Sciences
Quanzhou, China
xianweich@hotmail.com

Yongjun Zhou
Schoo1 of Naval Architecture, Ocean and Civil Engineering
Shanghai Jiao Tong University
Shanghai, China
yjzhou@sjtu.edu.cn

Abstract*—Chemical fiber paper tubes are the essential spinning equipment on filament high-speed spinning and winding machine of the chemical fiber industry. The precision of its application directly impacts on the formation of the silk, determines the cost of the spinning industry. Due to the accuracy of its application requirements, the paper tubes with defects must be detected and removed. Traditional industrial defect detection methods are usually carried out using the target operator's characteristics, only to obtain surface information, not only the detection efficiency and accuracy is difficult to improve, due to human judgment, it's difficult to give effective algorithm for some targets. And the existing learning algorithms are also difficult to use the deep features, so they can not get good results.Based on the Faster-RCNN method in depth learning, this paper extracts the deep features of the defective target by Convolutional Neural Network (CNN), which effectively solves the internal joint defects that the traditional algorithm can not effectively detect. As to the external joints and damaged flaws that the traditional algorithm can detect, this algorithm has better results, the experimental accuracy rate can be raised up to 98.00%. At the same time, it can be applied to a variety of lighting conditions, reducing the pretreatment steps and improving efficiency. The experimental results show that the method is effective and worthy of further research.

***Keywords*—Faster-RCNN; chemical fiber tube defect detection; convolutional neural network; region proposal network**

I. INTRODUCTION (*HEADING 1*)

In the modern chemical fiber industry, high-strength chemical fiber paper tube is a must for the production of accessories, its precision determines the production efficiency. For this reason, there is a high demand for the smoothness of the surface. However, in the process of the paper tube production, due to the production-level, the scene environment and other factors, it is easy to make the surface scratches, holes, joints and other defects. In the production, it is necessary to have a quick detection on the paper tube surface of the various defects, in order to control and improve the quality of the final product. Artificial visual sampling has been unable to meet the production of the number and speed requirements [1], so it is gradually developed into the computer software based online defect detection system. With the increasing detection quality requirements, as well as more sophisticated, faster production and other more stringent requirements of the process, the traditional machine vision-based detection method has been unable to meet the requirements of practical engineering.

There are several common types of defect detection algorithms: 1) Data-based algorithms, which are characterized by spatial distribution of image gray values in different ways, e.g. autocorrelation function, co-occurrence matrix [2], mathematical morphology [3], fractal dimension, etc.These methods use the surface gray value characteristics of the image, so it is simple and clear, the speed is fast for detecting the defects, but more susceptible to noise and external conditions of interference, and not easy to detect features with no obvious defects. 2) Spectral-based algorithm, through the expression of different features of the image, reflect the differences in the presence and absence of images in the image, e.g. Fourier transform, wavelet transform, Gabor transform [4], etc. This method can effectively detect the defects which are difficult to detect in the time domain by using the image frequency domain information, but there is a problem that both the whole and the local are difficult to be considered. At the same time, the wavelet basis of the wavelet transform is not easy to choose. If the detection is different, To ensure the detection accuracy, we need a lot of calculation and experiment. 3) Model-based algorithm, e.g. autoregression model [5], random field model [6], etc. By adjusting the parameters, we can determine whether there are defects, but if the parameters selected are improper, the convergence rate is slower. 4) Structure-based algorithm, using texture analysis method [7]. It has a good effect on the obvious defects of the texture, but it is difficult to

The work was partially supported by the National Natural Science Foundation of China (41274012)

2572-6609/17 $31.00 © 2017 IEEE
DOI 10.1109/ES.2017.35

choose the appropriate texture extraction algorithm for the defects which haven't obvious features. 5) Learning-based algorithms, e.g. BP neural network [8], support vector machines [9], etc. The feature information of the defective image can be further extracted, but the dimension of the information is still small, the precision is difficult to be improved and the false detection rate is also high. Usually a single method can not solve a specific problem, so it will be combined with a variety of methods to form a hybrid algorithm for testing.

In those traditional methods, most of the feature expressions are extracted from the manually designed features, there is an explicit preprocessing of the input signal, and the general ability is low. In addition, when the amount of data increases, the problem of overfitting will also be caught. For the internal joints used in the detection of paper tube, the characteristics are almost similar to the surrounding parts, sometimes even the human eyes are difficult to identify these defects, the use of surface features will inevitably lead to poor results. For the external joints, the forms are overabundant, the use of a single information detection method can not detect all the defects. With the development of deep learning in recent years, people have found a way to let the computer simulate the human brain to perceive the visual signal mechanism, and then design a deep network to achieve the visual function. CNN is a kind of deep neural network [10, 11], which uses the convolution structure to reduce the amount of memory occupied by the deep network, and can reduce the number of parameters of the network to alleviate the over-fitting problem of the model. Because CNN has these characteristics, Ross B. Girshick proposed RCNN method [12], obtained the candidate region by region selection method, and then input the region proposal to CNN for feature extraction. Finally, the feature of region extraction is classified and evolved into Faster-RCNN method used in this paper. By combining the CNN model with the Region Proposal Network (RPN), the target detection is achieved [13], and many better results are obtained when campared with the traditional method.

In this paper, a quick defect detection method is designed and implemented. For the training of the paper tube image collected by the line scanning camera this method using Fast-RCNN has the advantages of fast detection speed and good stability, and it is easy to configure the new defect type. It can greatly improve the production efficiency and reduce the manufacturing cost.

II. IMAGE ACQUISITION SYSTEM

At present, most of the printed defect detection systems use industrial computers, line scanning cameras, rotary encoders and linear light source as the solution. At the time of production, the linear condensing cold light source composed of bright LEDs is irradiated on the surface of the paper tube by means of transmission or reflection, and is triggered by a rotary encoder synchronized with the paper tube operation, so that the line scanning camera mounted on the production line is scanned synchronously, then transmit the paper tube images collected by the camera through the industrial computer acquisition card to the image processing system software for defect identification processing. Paper tube defect detection system diagram is shown in *Fig. 1*. Because of the obvious difference between the feature extracted by the defective image after the deep neural network and the characteristics of the normal image, it can detect the defect by identifying the feature, then judge, classify and follow the defect at the same time.

During the real production, with the higher precision requirements and the faster production speed, as well as the larger amount of image data collected in the unit time, the current image processing algorithms can not meet the requirements more, and the increase in the number of defects will also lead to an increase in the amount of calculation, thereby increasing the development costs. So we use the Faster-RCNN method, with a high detection speed and accuracy at the same time, for different defects, we do not need to carry out additional development, only need to use these additional defects to further train the model, so that it can meet the industry production requirements better.

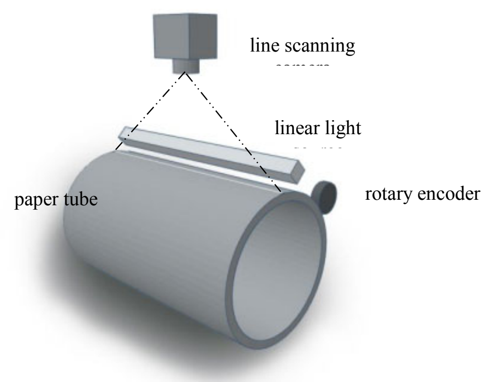

Fig. 1. Image Acquisition System

III. THE DEFECT DETECTION METHOD

BASED ON FASTER-RCNN ALGORITHM

At present, the defect detection tasks in industrial production generally follow the traditional methods such as covariance matrix, support vector machine and wavelet transform. These methods have proven their effectiveness in long-term applications, but still have a lot of breakthrough space, there are several shortcomings:

1) Different defects vary widely, so we need targeted, different detection methods.

2) For the defects that is difficult for naked eye to identify, we can hardly find targeted identification method through deduction.

3) The image acquisition condition is required to be high, and the defect images acquired under difficult conditions can not be recognized.

In view of these above problems, this paper uses the Faster-RCNN method which can automatically extract a large number of features and can detect each kind of object to extract and identify the image target area.

The traditional method of image detection is the Deformable Parts Models [15] (DPM), in the VOC2007 data set it can achieve 43% of mean Average Precision [16] (mAP). With the development of deep learning in recent years, Ross B. Girshick has implemented the CNN powerful classification ability to apply to the image detection, developed the RCNN series image detection algorithm. The latest Faster-RCNN method can achieve 73% mAP on VOC2007. Its excellent results show that this method in the face of various categories in a variety of background has a strong ability of distinguishing, and the defect detection in the industrial production has a more simple background than the VOC2007 data set, so using this method can achieve better results.

A. Region Proposal Network

As shown in *Fig. 2*, the center point of the sliding window on the feature map can be mapped back to the input image, corresponding to a center point position in the input image. In order to make the algorithm better applied to different shapes and sizes of the target, we set multiple anchor points to the same position on the feature map to predict the bounding box of different scales and different ratio aspect of the input image. Usually the anchors correspond to the input image of the three scales and three kinds of ratio aspects of the candidate area, which produces nine anchor points. This means that nine regions can be generated for each region. After scanning, a total of 9*H*W*D candidate regions are predicted for the feature graph of the model output, where H and W are the height and width of the feature map, and D means the dimension of the feature map.

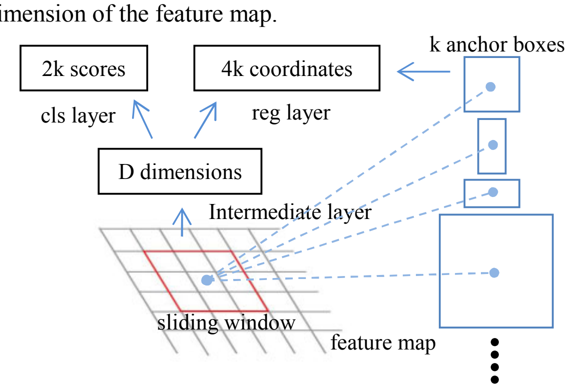

Fig. 2. The Anchors

B. The Struct of Faster-RCNN [13]

Faster-RCNN network structure is shown in *Fig. 3*:

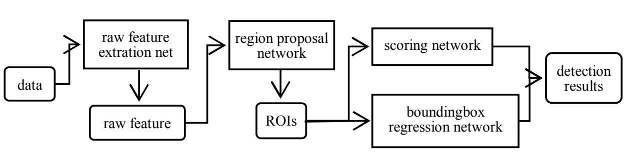

Fig. 3. The Struct of Faster-RCNN

After the data is input, it is send to a feature extraction network, usually it is a 5 shared layer ZF net (output 256-dimisional feature map) or a 13 shared layer VGG net (output 512-dimisional feature map). We use ZF network as an example in this paper, get a total 256 channels, for the feature map of H*W size, put the feature map through a *conv* layer and a *relu* layer into the region proposal network. In the region proposal network, we use multi-scale anchors to extract the target candidate area (Region Proposal). Use the convolution kernel (sliding window) on the feature map extracted by the network to convolute with the feature map, obtain a 256-dimensional feature vector. Then access to the two fully connected layer, determine the category and regress the bounding box for the region corresponds to the anchor in the origin graph. The category determination layer outputs two parameters for estimating whether the original map area belongs to the target or not. The bounding box regression layer contains four coordinate elements that represent the amount of modification of the bounding box on the original area relative to the anchor point, it is used to determine the exact location of the target. After scoring and bounding box regression for each anchor, the bounding boxes are subjected to non-maximum suppression, and then the close bounding boxes are merged to finally get the target object to be detected.

IV. EXPERIMENTAL DESIGN AND RESULTS

A. Chemical Fiber Paper Tube Defect Database

In the chemical fiber paper tube production process, will produce stains, machine scratches, production process problem and other issues. These problems will lead to defects, the main types of defects are internal joints, external joints, scratches and so on. The defective image is shown in Figure 4 Inside the red box:

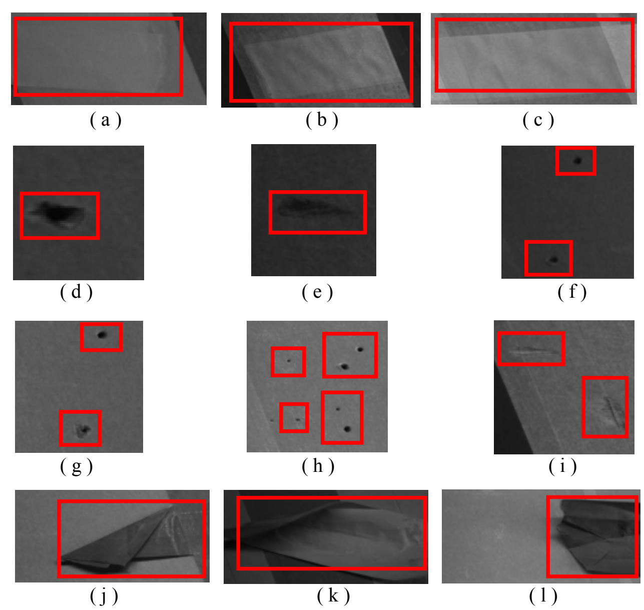

Fig. 4. Defects of All Kinds (a) (b) (c) are internal joints (d) (e) (f) (g) (h) (i) are scratches (j) (k) (l) are external joints

Fig. 4(a) to *Fig. 4(c)* are internal joint images, in the paper drum production, the paper exhausted, so the workers paste a scrip as an identification, then the entire scrip was inserted into the paper, form the inner joint. As the joints are thin, it is easy to confuse of the appearance, not only difficulty for the traditional method to identify, some internal joints are even difficulty for human eyes to identify. *Fig. 4(d)* to *Fig. 4(i)* are scratch images, mostly due to the machine scratching on the paper drum surface. The scratches are usually small, traditional methods are difficult to guarantee the recognition rate. *Fig. 4(d)* to *Fig. 4(l)* are the external joint images, the reason of the formation is similar to the internal joints, but parts of the joints are exposed outside the roll paper. This kind of defects usually occur with the inner joints.

A total of 132 internal joint images, 92 external joint images and 133 scratch images were used in this paper. The testing samples were randomly selected 50 per time for 10 fold cross validation, the internal joints and scratch images were trained using 15, 35, 50, 75 pictures, the external joint images were trained using 15, 25, 35, 40 pictures. The results of testing were finally compared with the results obtained using the traditional method.

B. *Experimental Results*

In this paper, the three kinds of defect data described in 4.1 were trained and tested by Faster-RCNN. The VGG-16 [17] net using 13 shared layers and the ZF [18] net using 5 shared layer were used respectively, and the pre-training was based on ImageNet image library, then we used the defect images for fine tunning to get the model to test the testing images. At the same time, we used the traditional methods which have achieved good results in image recognition quests, these were SIFT pattern matching method [19] and linear SVM classifier, we tested the same testing images for comparison. The SIFT pattern matching algorithm used 75 patterns for detecting. The SVM method divided the images into small pieces, used the texture feature and the gray histogram of each small blocks to train the classifier, then we used the classifier to judge whether each block of the images included the defects.

During the experiment, the average training time of Faster-RCNN with ZF was about 12.5 minutes per image, and the average training time of Faster-RCNN with VGG-16 was about 22 minutes per image, the average SVM method training time is about 0.39s per image. During the test, the model trained using the Faster-RCNN method was approximately 0.13 seconds per image, the average testing time of the SIFT pattern matching method was 2.8 seconds per image, and the SVM method detected an image for 0.015 seconds. When the intersection-over-union (IoU) is higher than 0.7, the final defect image detection results are as follows:

Internal joints: The results of the internal joint detection are shown in Table 1:

TABLE I. INTERNAL JOINT DETECTION RESULTS

Methods	Tarining Set Amount	Precision	Recall
Faster-RCNN + ZF	15pics	47.0%	92.2%
	25pics	54.6%	95.8%
	50pics	60.2%	97.7%
	75pics	90.8%	99.3%
Faster-RCNN + VGA	25pics	48.2%	96.2%
	75pics	93.2%	99.2%
SIFT Pattern Matching	75 pics	6.6%	10.6%
SVM Feature Classifier	75 pics	53.2%	40.7%

It can be seen from Table 1 that the SIFT pattern matching method used to detect the internal joints, only obtained the detection rate of 6.6%, because the feature points of the internal joint images were not obvious, and the recall rate was 10.6%, which was difficult to be applied to industrial production. Using the SVM method for classification increased the detection rate to 53.2% and the recall rate to 40.7%, this meant that the number of false retrieval increased sharply due to the insignificant feature of the internal joint, so that this method is difficult to use in industrial production. When using the Faster-RCNN method, as the number of training images increased, the detection rate increased, the number of truly detected defects increased, as well as the recall rate increased, which means that the number of detection errors was reduced. When using 75 training images, the detection achieved high accuracy. At the same time because of the increase in the number of layers, we can see that the use of VGG-16 net could help to extract deeper levels of features, so the training accuracy increased a bit, but the training time increased almost one time than the ZF net. Therefore, when the Faster-RCNN method was used for the detection of the internal joint, the detection result is considerably improved compared with the traditional methods.

External joints: The results of the external joint images are shown in Table 2:

TABLE II. EXTERNAL JOINT DETECTION RESULTS

Methods	Tarining Set Amount	Precision	Recall
Faster-RCNN + ZF	15pics	63.2%	94.0%
	25pics	74.6%	99.7%
	50pics	80.4%	100%
	75pics	97.8%	100%
Faster-RCNN + VGA	25pics	64.8%	98.3%
	75pics	98.0%	100%
SIFT Pattern Matching	75 pics	25.0%	25.6%
SVM Feature Classifier	75 pics	79.2%	85.3%

As we can see from Table 2, in the detection of external joints, the use of SIFT pattern matching method only achieved 25% detection rate because the external connector image types are not easy to match and achieved a recall rate of 25.6%. The use of SVM method achieved the detection rate of 79.2%, but the recall rate rose to 85.3%. When using the Faster-RCNN method, the detection rate and the recall rate increased when the number of training images increased, the same as the inner joint images. Compared with the internal joints, because the deep features of the external joints were more obvious, all the detection accuracy of the methods had increased. And because of the high separability, the recall rate in the test can reach to 100%. It can be seen that because of the obvious characteristics and the strong separability of the internal joints, we got better detection results.

Scratches: The results of the scratch images are shown in Table 3:

TABLE III. SCRATCH DETECTION RESULTS

Methods	Tarining Set Amount	Precision	Recall
Faster-RCNN + ZF	15pics	61.4%	98.1%
	25pics	76.0%	99.2%
	50pics	83.0%	100%
	75pics	95.6%	100%
Faster-RCNN + VGA	25pics	77.2%	100%
	75pics	97.6%	100%
SIFT Pattern Matching	75 pics	48.0%	38.7%
SVM Feature Classifier	75 pics	92.2%	83.5%

We can see from Table 3, in the detection of scratches, because the joints were too small, it was difficult to match the feature points, the use of SIFT pattern matching method had low detection rate and recall rate. The SVM method used the block discrimination mode, which has 92.2% detection rate and 83.5% recall rate for this kind of small defects. When using the Faster-RCNN method, the detection rate increased with the increasing of the number of training samples, and the test result could reach 95.6% when using 75 samples. At the same time, due to the scratches had more similar characteristics the recall rate was close to 100%.

V. CONCLUSION

In the industrial production of paper drum defect detection, the efficiency and accuracy of detection are required. The use of traditional methods have a large gap of the recognition accuracy between different problems. In practical applications, the traditional methods can also hardly achieve real-time detection. Faster-RCNN solves the problem of extracting shallow features by means of CNNs. At the same time, the bounding box selection and category scoring are carried out through the region proposal network, no extra pre-processing is needed, with end-to-end training training methods. Through this method, we first get a higher accuracy rate, and because of the integration of the bounding box extraction process, the detection time is shorter. Meanwhile, we can use the newly labeled data continuous train the model to further improve the accuracy rate, the Faster-RCNN method applied to the paper drum defect detection of industrial production will achieve a very good effect.

However, Faster-RCNN method also has somgthing insufficient that need to be improved in future studies. First of all, although it has a good accuracy during experiments, because we need to detect the defects in paper tube production as many as possible, there is still room for improvement in recognition rate. And second, the training time is too long, resulting in a long time preparation in advance. Finally, in the bounding box regression the method may select more parts of the object, we can improve it to be more intuitive in the application.

REFERENCES

[1] J. Leel, P. Park. Inspection of defect on LCD panel using local mean algorithm based on similarity[J]. 13th International Conference on Control, Automation and Systems (ICCAS)，2013：584-588.

[2] 于国防. 煤矸区分中的间隔灰度压缩扩阶共生矩阵[J]. 中国图象图形学报, 2012, 17(8): 966-970.

[3] Liu X, Wen Z, Su Z. Choi. Slub extraction in woven fabric images using Gabor filters[J]. Textile Research Journal, 2008, 78 (4): 320-325.

[4] 林如意. 基于图像处理技术的纺织品瑕疵检测方法[D]: [硕士学位论文]. 杭州：杭州电子科技大学, 2013.

[5] Haralick R M, Statistical and structural approaches to texture[C]. Proceedings of the IEEE, 1979, 67(5): 786-804.

[6] Serafim A F L, Multiresolution Pyramids for Segmentation of Natural Images Based on Autoregressive Models. Application to Calf Leather Classification[C]. Proc.IEEE Int'l Conf. Industrial Electronics, Control & Instrumentation (IECON91), 1991, (3): 1842–1847.

[7] 祝双武, 郝重阳, 李鹏阳等. 基于纹理结构分析的织物疵点检测方法[J]. 计算机应用, 2008, 28(3):647-649.

[8] Stojanovic R, Mitropulos P, Koulamas C. Real-time vision-based system for textile fabric inspection [J]. Real-time Imaging, 2001, 7: 507-518.

[9] 丁世飞, 齐炳娟, 谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 2011, 40(1): 2-10.

[10] D. Eigen, J. Rolfe, R. Fergus. Understanding deep architectures using a recursive convolutional network[R]. arXiv:1312.1847v2, 2014.

[11] K. Simonyan, A. Zisserman.Very deep convolutional networks for large-scale image recognition[R]. In arXiv:1409.1556, Sept. 4, 2014.

[12] Girshick R, Donahue J, Darrell T, et al. Region-Based Convolutional Networks for Accurate Object Detection and Segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(1):142-158.

[13] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networs [C]// Advances in Neural Information Processing Systems, 2015: 91-99.

[14] 韦玮. 基于机器视觉的防伪纸安全线检测关键技术研究与应用[D].机械科学研究总院,2007.

[15] Felzenszwalb P F, Girshick R B, Mcallester D, et al. Object detection with discriminatively trained part-based models.[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2010, 32(9):1627.

[16] Ceri S, Bozzon A, Brambilla M, et al. An Introduction to Information Retrieval[J]. Pharmacogenomics Journal, 2013, 2(2):96-102.

[17] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Eprint Arxiv, 2014.

[18] Zeiler M D, Fergus R. Visualizing and Understanding Convolutional Networks[J]. 2013, 8689:818-833.

[19] RACHEL L T, SIEBERT J P. Local feature extraction and matching on range images: 2.5D SIFT [J]. Computer Vision and Image Understanding, 2009, 113 (12): 1235-1250.