Using Faster-R-CNN to Improve the Detection Efficiency of Workpiece Irregular Defects[1]

Zhao Liu, Yan Li (Corresponding author)
Dept. of Computer Science and Engineering, Inha University

Abstract

In the construction and development of modern industrial production technology, the traditional technology management mode is faced with many problems such as low qualification rates and high application costs. In the research, an improved workpiece defect detection method based on deep learning is proposed, which can control the application cost and improve the detection efficiency of irregular defects. Based on the research of the current situation of deep learning applications, this paper uses the improved Faster R-CNN network structure model as the core detection algorithm to automatically locate and classify the defect areas of the workpiece. Firstly, the robustness of the model was improved by appropriately changing the depth and the number of channels of the backbone network, and the hyperparameters of the improved model were adjusted. Then the deformable convolution is added to improve the detection ability of irregular defects. The final experimental results show that this method's average detection accuracy (mAP) is 4.5% higher than that of other methods. The model with anchor size and aspect ratio (65,129,257,519) and (0.2,0.5,1,1) has the highest defect recognition rate, and the detection accuracy reaches 93.88%.

I. Introduction

After entering the era of big data, higher requirements have been put forward for product defect detection in industrial production. Advanced detection technology algorithms are widely used in industrial irregular defect detection. Based on understanding the problem of industrial product defect detection, this paper analyzes the irregular defect detection technology with a depth learning algorithm as the core. This paper studies how to properly adjust the structure and parameters of the backbone network when improving the deep learning Faster R-CNN network model, and adds deformable convolution to enhance the effect of feature extraction. According to the collected industrial data set characteristics, the image processing technology is used for optimization processing, and finally, the super parameter setting of the optimal model is obtained in the experimental analysis to improve the application performance of the model. At the same time, according to the workpiece surface defect detection, the depth learning network algorithm can be reasonably used to obtain the defect information in the collected workpiece image, and quickly learn the surface defect characteristics, to build a network detection and recognition model.

II. Related research

This paper proposes a detection algorithm based on the Faster R-CNN network structure model, which is mainly used to automatically locate and classify the defect areas of the workpiece. As an image detection method, CNN (Convolutional Neural Network) can enhance the detection accuracy by adding deformable convolution to enhance feature extraction and obtain target defects based on strengthening practical operation. After training, the Faster R-CNN network structure model will find the defect area from the detection image and mark it effectively.

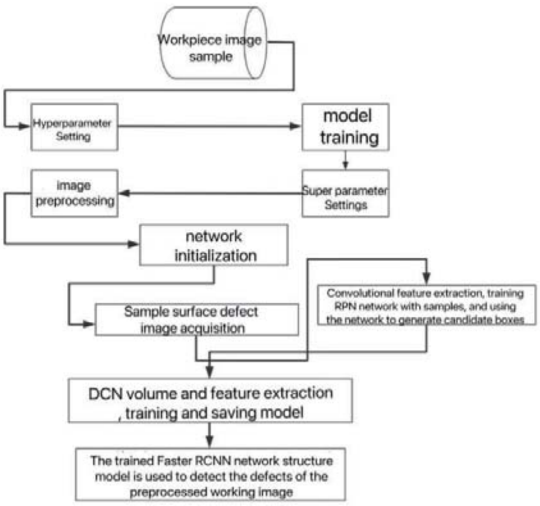

Figure 1. Automatic detection process of workpiece defects

2.1 Improved Faster R-CNN algorithm and design

According to the analysis of the Faster R-CNN structure shown in Figure 2, the overall algorithm includes two major modules, one is the candidate frame extraction module of RPN (area generation network), and the other is the Faster R-CNN detection module. Among them, Fast R-CNN can use CNN to obtain feature map in RPN. Rol In the pooling layer, the Rol maximum pool is selected for feature mapping, and a fixed-size feature vector is extracted from each Rol. For all Rolls, the four parameters of the regression output represent the central coordinates of the object's bounding box(T_x^μ, T_y^μ), height$T_h^\mu T_w^\mu$, softmaxOutput the probability of k+1 class (k training class+1 background class)$P = (P_0, P_1, \dots, P_k)$.If the IoU ratio between Rol and the real

1 Acknowledgement: This research was supported by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (P0017157-World Class Plus Project)

frame exceeds 0.5, the actual label is positive; On the contrary, if the IoU ratio value is controlled between 0.1 and 0.5, then the label is the background. Use the gradient descent formula of the too small batch as the loss function to train the end-to-end Fast R-CNN network model, where u represents the boundary box labels of all real objects, V represents the coordinates of the bounding box of all real objects, $v = (v_x, v_y, v_h)$.The parameter u is 1 or 0, which represents the Rolls of the target or background, and La represents the logarithmic loss and SoftMax loss, L_{reg}1 Used to regress losses. The specific formula is as follows:

$$L(P, u, T^{\mu}, v) = L_{cls}(P, u) + [u \geq 1] \sum_{i \in \{x,y,w,h\}} L_{cls}(T^{\mu}, v_i) \quad (1)$$

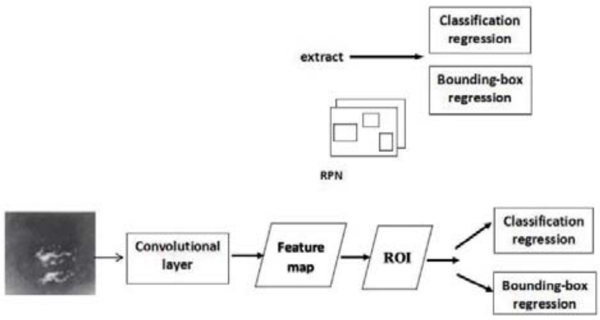

Figure 2. Faster R-CNN structure

The backbone network of the deep learning network studied in this paper is ResNet101. The maximum number of channels in the convolution layer can reach 23. The small convolution kernel is more nonlinear than the large convolution kernel; the ResNet101 network model is composed of L residual modules, The input and output are xl, xl+ 1, and the calculation formula is as follows:

$$y_L = h(x_L) + F(x_L, w_L) = w_s x_L + F(x_L, w_L) \quad (2)$$

$$x_{L+1} = f(yL) \quad (3)$$

$$Assume\ x\ first\ \ (L+1) = f(yL), w\ \ S = I, where\ \ w\ \ S\ is\ the\ identity\ matrix, h\ (xL) = xL, in\ summary \quad (4)$$

$$x_L = x_{L-1} + F(x_{L-1},\ w_{L-1}) = x_0 + \sum_{i=0}^{L-1} F\,(x_i, w_i) = x_l + \sum_{i=1}^{L-1} F\,(x_i, w_i) \quad (5)$$

2.2 ResNet101 Improved model

The backbone network of the Faster R-CNN network model studied in this paper is improved based on ResNet101 (figure 3). The improvement means continuously expanding the number of channels in the original network model. Four channels are added to the third layer of the convolution network, and 13 channels are added to the fourth layer of the convolution network. Finally, the fourth field network is transplanted to the fifth layer, and four channels are added. This improvement is because it only deepens the network depth, which will reduce the detection effect. From the improved network model; First, we need to input the feature image and use the sliding window to obtain the anchor point; Second, after the introduction of deformable convolution, the original convolution network is divided into two parts. One part uses additional convolution layers to learn the target offset, outputs to obtain the offset in two directions, and uses the offset window to replace the original sliding window, and then ensures that the calculation process is consistent with the conventional convolution; Third, the network algorithm has defects in identifying irregular targets, which stems from its own network structure; Fourth, since different areas may correspond to objects of different sizes, in order to improve the accuracy of defect detection, the scale or perception size should be automatically adjusted. In this process, the benchmark network is used to learn the offset to ensure that the convolution kernel is offset at the sampling point, focusing on the target area.

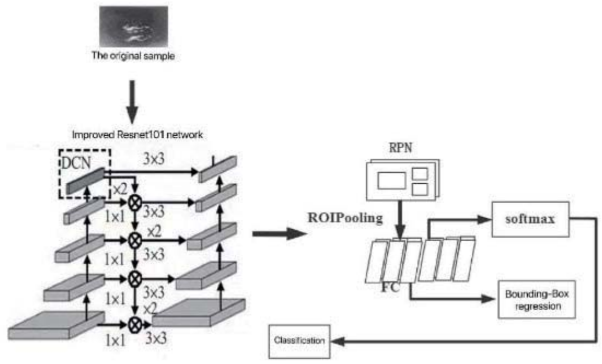

Figure3. Faster R-CNN Model Diagram

III. Experiment

In this paper, the metal workpiece is selected as the sample, and the number of images in the experiment is 3566, which mainly includes five defects: jet flow/pitting/ bottom leakage/mottled/paint bubble. A labeling annotation tool was used to effectively label the defect types and boundary box coordinates in the image, and 5684 target objects were finally found, The efficiency of defect detection can be further improved by transforming the labeled image into a VOC2007 data format. In the experimental analysis, 20% of the images shall be selected as the test set, and the remaining 80% shall be the training set and verification set. Before the formal operation analysis, the method of data expansion is used for optimization processing to ensure the accuracy of defect detection.

In the field of target detection, mAP (mean Average Precision) can calculate the average value of each category. IOU (Intersection Over Union) can be divided into foreground and background according to the numerical value. Generally, a value greater than or equal to 0.7 is the foreground, and a value lower than 0.3 is the background. It was found in the experiment that a higher mAP value could be obtained in the analysis by properly adjusting this value.

$$AP = \int_0^1 p(r)dr \quad (6)$$

P represents the accuracy rate, and r represents the recall rate; Accurate calculation of AP is to analyze the area below the P-R curve.

$$\mathrm{MAP} = \frac{1}{\mathrm{N}}\sum \mathrm{AP_i} \tag{7}$$

In the above formula, N represents the number of categories; MAP belongs to the average value of AP value under all categories, and the calculation formulas of accuracy and recall are as follows:

$$\mathrm{Accuracy} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{8}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP + FN}} \tag{9}$$

TP (True Positive) represents the sample data set whose prediction result is positive; FP (False Positive) represents the sample data set whose negative sample prediction result is positive; FN (False Negative) represents the sample data set with negative prediction results. In this experimental study, we should focus on the AP numerical results of the model, If there is a category in which the classification effect is poor, the application model should be adjusted to optimize the detection effect.

In this research experiment, we use TensorFlow deep learning framework to build a network model. When training the RPN and Fast R-CNN network models, the average value of the parameters of the full connection layer should be considered as 0, the standard deviation is 0.001, the Gaussian distribution is initialized, the learning rate is 0.001, the momentum is 0.9, and the attenuation factor is 0.0005.

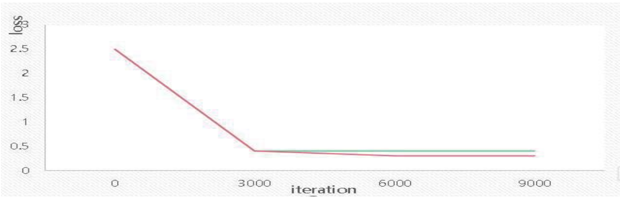

Figure 4. Influence of IoU value on loss curve

3.1 Experimental result

During the training, 5 groups of point combination models with different width-height ratios and different scales were selected, and the corresponding features were obtained for training tests. When the anchor size is (65129257519) and the aspect ratio is (0.2, 0.5, 1, 1), the model has the highest recognition rate of defects. The final results show that the average detection efficiency of the model can reach 93.88% when resolving the five defects proposed in this study. Combined with the maximum intersection and union ratio analysis of figure 4, it can effectively change the convergence effect of the loss curve, and the effect is strongest when the value is 0.9 and 0.1.

IV. Conclusion

Based on the Faster R-CNN network model, this paper proposes a metal workpiece surface defect detection system with depth learning as the core. Due to the inherent defects of the traditional depth learning algorithm, the recognition effect on the workpiece surface defect detection is poor. Therefore, based on integrating the previously accumulated experience, the convolution network is used to automatically extract features for detection, By using the ResNet101 model for migration learning, the research on irregular defect detection of workpieces is strengthened, and the application quality of industrial products is guaranteed on the basis. According to the experimental results, the model with the anchor size and width height ratio of (65129257519), (0.2,0.5,1,1, 1) respectively has the highest recognition rate of defects, and the detection accuracy reaches 93.88%. Due to the deep network structure of the feature extraction network ResNet101, the average detection speed of each image is slightly slower than that of the Faster R-CNN network model, which needs to be further improved.

References

[1] Cnuz A, Simas F.Efficient feature selection for neural network based detection of flaws in steel welded joints using ultrasound testing [J]. Ultrasonics,2016.

[2] UIJLINGSJR R,SANDEK EA V D,GEVERS T, et al.Selective Search for Object Recognition[J]. In-ternational Journal of Computer Vision, 2013, 104(2):154-171.

[3] LONG J,SHELHAMER E, DARRELLT.Fully Convolutional Networks for Semantic Segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2015:3431-3440.

[4] Rezatofighi H, Tsoi N,Gwak J Y,et al.Generalized intersection over union :a metric and a loss for bounding box re-gression[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition ,2019:658 - 666.[17]