G. Z. OUYANG, W. Y. ZHANG

ISSN 0543-5846
METABK 62(1) 61-64 (2023)
UDC – UDK 669.14-621.41:536.5-621.771:621.9.015.004.64-620.19:62.001.57=111

A SURFACE DEFECT DETECTION METHOD OF STEEL PLATE BASED ON YOLOV3

Received – Primljeno: 2022-05-31
Accepted – Prihvaćeno: 2022-08-10
Original Scientific Paper – Izvorni znanstveni rad

At present, the steel plate surface defect detection technology based on machine vision and convolutional neural network (CNN) has achieved good results. However, these models are mostly two-stage methods, extracting features first and then classifying them, which is slow and inaccurate. Therefore, this paper proposes a single-stage surface defect detection method of steel plate based on yolov3, which can classify defects, determine the location of defects, and greatly improve the detection speed. It is of great significance to realize the automation of cold rolling production line. The experiment shows that the detection speed of this model reaches 62 fps and the accuracy reaches 73 %, which has a good prospect in industry.

Keywords: steel plate, cold rolling, surface defect detection, yolov3, deep learning

INTRODUCTION

The surface quality of steel plate is one of the most important quality factors of steel plate. When there are defects on the surface of the steel plate, it will reduce the performance of the product and cause huge economic loss. Therefore, it is of great practical significance to find defects quickly and accurately, clarify the category and location of defects, and improve the cold rolling processing of steel plate according to the causes of defects in time.

The previous detection methods of steel plate surface defects, such as manual visual sampling method, eddy current detection technology, infrared detection technology, magnetic flux leakage detection technology and machine vision technology, are easily affected by subjective factors, requiring to artificially set the algorithm to extract features. The poor detection speed and effectiveness are not suitable for industrial production requirements.

With the emergence of AlexNet [1], the deep learning (DL) technology represented by convolutional neural network is widely used in steel plate surface defect detection because of its self-learning ability, which can automatically complete feature extraction by simply inputting the defect images collected into the network model. In [2] a five-layer convolutional neural network is used to classify the surface defects of hot-rolled steel plates. In [3] an improved mask R-CNN algorithm is proposed to improve the accuracy, but the speed of detection only reaches 5,9 fps. In [4] a multi-level feature fusion convolutional neural network based on ResNet is proposed, and the detection accuracy is up to 82,3 % and the speed is up to 20 fps. In [5] a surface defect detection method of steel plate based on Faster-RCNN is proposed, with a speed of 21,42 fps and a classification accuracy of 87,14 %. However, a series of networks such as Faster R-CNN are two-stage detection networks, which are divided into two steps: candidate region extraction and classification. The speed of defect detection is slow, and the detection accuracy of classification and location needs to be further improved.

In this paper, a steel plate surface defect detection model based on YOLOV3 [6] is proposed. This model has the characteristics of single-stage detection, determining the location of defects and good detection effect. The regression idea is used to directly predict the location and category of defects, which largely makes up for the region-based algorithm and greatly improved the speed. The location of defects also provides a basis for the subsequent optimization of the cold rolling processes. The results verify the effectiveness of the model, which is able to detect defects quickly and accurately.

PROCESS OF COLD ROLLED STEEL PLATE

The key processes of cold rolling include pickling, cold rolling, annealing and leveling. The function of pickling process is to remove oxide scale from the surface of hot rolled plate for cold rolling process. Cold rolling is the key process in the production of cold rolled plate. The function of the annealing process is to improve the plasticity of cold-rolled plates and to restore the plates to good machining performance as much as possible. The function of leveling process is to improve the flatness of the plate, and achieve a certain roughness on the surface of the plate. The process flow of cold rolled steel plate is shown in Figure 1.

G. Z. OuYang, W. Y. Zhang, School of Computer Science and Software Engineering, University of Science and Technology Liaoning, China. Corresponding author: W. Y. Zhang (zhangwenyu8518@126.com)

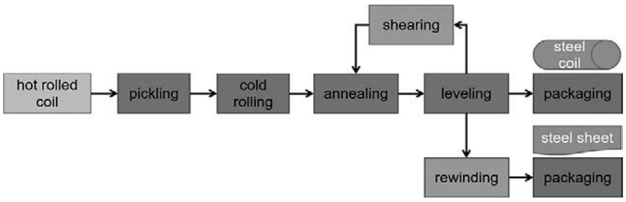


Figure 1 The process of cold rolled steel plate

CAUSE ANALYSIS OF SURFACE DEFECT OF STEEL PLATE

There are many processes in the production of cold rolled steel plate, and the surface defects of steel plate inevitably appear frequently. In order to ensure the stable operation of the cold rolling line, these defects must be detected in time, the common defects are shown in Figure 2.

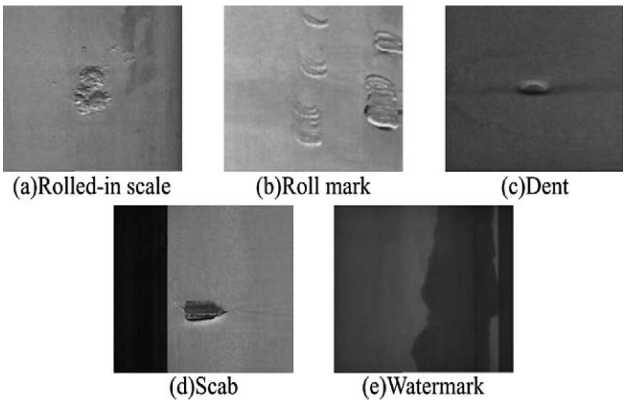


Figure 2 Surface defects image of steel plate

Rolled-in scale: as shown in Figure 2 (a). A spot or fish scale defect of varying sizes is formed on the surface of the steel plate. The causes: the iron scale of slab cannot be completely removed during the heating process; The failure of phosphorus removal equipment.

Roll mark: as shown in Figure 2 (b). Periodic pits or bumps appear on the surface of steel plate. The causes: the strip weld is too high or unevenly cleaned; Hard metal or dirt stuck on the roller leaving indentations.

Dent: as shown in Figure 2 (c). On the surface of the steel plate, one side depression, the other side corresponding to the same position. The causes: something adheres to the roller; Something adheres to the steel strip.

Scab: as shown in Figure 2 (d). Irregular strips of upturned metal layer appear on the surface of the steel plate. The causes: the residual scars on the surface of the slab are not cleaned, and they are left on the steel plate after rolling.

Watermark: as shown in Figure 2 (e). The surface of the steel plate is attached with water-like ripples. The main reason for its formation is the lack of lubrication and extreme pressure of emulsion. This problem can be improved by adjusting the injection angle of emulsion nozzle and increasing the surface roughness of the working rollers.

The above are the common defects and causes of steel plate. Since most of the defects generated during the production process of this steel mill are watermarks, and the others are few defects. In order to balance the data, we group the remaining four defects into one category, which is called marking. If defects are detected, the corresponding improvement methods can be used according to the category of defects.

SURFACE DEFECT DETECTION PROCESS OF STEEL PLATE

The traditional steel plate surface defect detection process is shown in Figure 3 (a). The system acquires the images of the plate on the production line in time through the camera. After the operations of pre-processing, segmentation and feature extraction, the extracted features are put into the classifier for training. Finally, get the classification results of the defect detection.

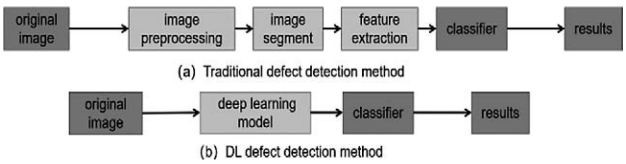


Figure 3 Comparison of the traditional defect detection method and the DL defect detection method

The process of steel plate surface defect detection based on deep learning network is shown in Figure 3 (b). Compared with the traditional detection methods, deep learning networks do not need to separate the two processes of feature extraction and classification training. It can automatically extract the better features and achieve accurate detection results faster.

It can be seen from the Figure 3 that in order to meet the requirements of defect detection timeliness, it is necessary to take the time from the input of the image to the classification result less than the time of the next image acquisition, otherwise there will be missed detection on the surface of the steel plate, which puts forward higher requirements for rapid and efficient defect detection method.

STRUCTURE OF THE MODEL

The model is proposed based on yolov3, which is different from Faster-RCNN and other networks. It transforms the target detection problem into a logistic regression problem, and divides the image into several grid cells. Each grid cell is responsible for detecting the corresponding defects, and then the localization and classification results of defects are directly output.

This model adopts Darknet-53 as the backbone network of feature extraction, which is composed of 53 convolution layers. Each convolution layer is composed of convolution operation, normalization and activation function. Moreover, the ResNet [7] is added to set up a fast link between layers, which can deepen the network depth and improve the detection speed of the model. The 1 × 1 convolution in the residual network can save resources and accelerate the training of the network.

Due to the more information loss over convolution, the model also uses the method of feature pyramid fusion, which fuses 8 times, 16 times and 32 times down-sampling feature maps respectively to prevent excessive information loss and improve the detection effect of defects. The multi-scale detection method is used in the head network, to detect the feature maps of three down-sampling sizes of 13 × 13, 26 × 26, 52 × 52. The feature maps of each size are set with 3 anchor boxes, with a total of 9 anchor boxes, to detect the large, medium and small defects respectively. Because the deeper the down-sampling is, the smaller the size of the feature map is, and the larger the size of the original map represented. Therefore, three large anchor boxes are used to detect large defects on the smallest feature map. Three medium-size anchor boxes are used to detect medium-size defects on medium-size feature maps. On the largest feature map, three small boxes are used to detect small defects. Each grid cell only uses three anchor boxes to detect corresponding defects, which makes the model detect faster while maintaining accuracy. The network structure of the model is shown in Figure 4.

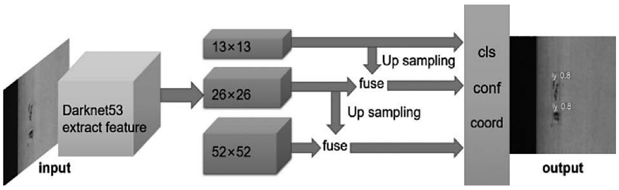


Figure 4 The network structure of this model

The loss function of this model is shown in Formula (1), which consists of three parts: confidence error (IouErr), classification error (ClsErr) and position error (CoordErr), which are Formula (2), (3) and (4), respectively.

$$Loss = \sum_{i=0}^{s^2} iouErr + clsErr + coordErr \qquad (1)$$

$$\sum_{i=0}^{s^2}\sum_{j=0}^{B} L_{ij}^{obj}(C_i - \hat{C}_i)^2 + \lambda_{noobj}\sum_{i=0}^{s^2}\sum_{j=0}^{B} L_{ij}^{noobj}(C_i - \hat{C}_i)^2 \qquad (2)$$

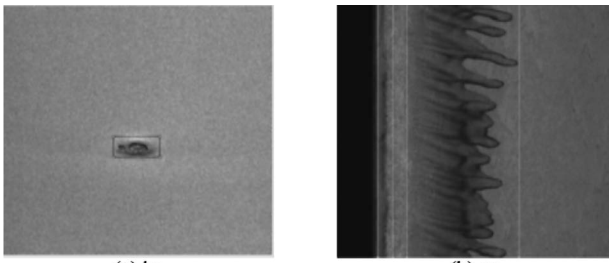
$$\sum_{i=0}^{s^2} L_i^{obj} \sum_{j=0}^{B}[(P_i(c) - \hat{P}_i(c))^2] \qquad (3)$$

$$g\lambda_{coord}\sum_{i=0}^{s^2}\sum_{j=0}^{B} L_{ij}^{obj}[(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2]$$
$$+ g\lambda_{coord}\sum_{i=0}^{s^2}\sum_{j=0}^{B} L_{ij}^{obj}[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \qquad (4)$$

Where C_i is the confidence score, $\hat{C}_1$ is the IOU (intersection over union) between the ground truth and the bounding box, w_i, h_i, x_i and y_i are the width, height and center coordinates of the bounding box, $\hat{w}_1$, $\hat{h}_1$, $\hat{x}_1$ and $\hat{y}_1$ are the width, height and center coordinates of the ground truth. L_{ij}^{obj} indicates whether there is a defect in the grid cell. If there is a defect, the value is one; If there is no defect, the value is zero. When it is zero, the classification task is not performed. λ_{noobj} and λ_{coord} are the weight coefficients. It can be seen from the formula that the model outputs classification results and position information at the same time, which not only ensures the detection effect, but also improves the detection speed.

EXPERIMENT AND ANALYSIS

The experimental data set was collected from the cold rolling mill by CCD camera online image acquisition technology. There are 2 653 defect samples in total. We divide them into watermark and marking, and there are 1 175 defect samples of marking and 1 478 defect samples of watermark. Use 70 % of the dataset as training samples and 30 % as test samples. In these experiment samples, each image contains at least one defect. The defect samples in this paper are marked with rectangular boxes by Labelimg. The ly represents the defect of marking, and its annotation is shown in Figure 5 (a); The sy represents the defect of watermark, and its annotation is shown in Figure 5 (b).



(a) ly (b) sy

Figure 5 Annotation sample by Labelimg

In the target detection, the average accuracy AP (Average Precision) is calculated by using the P (Precision) and R (Recall) of a certain category of model. Finally, the mean average precision (mAP) of all classes is used as the performance evaluation standard of the target detection model. Using the following formula (6):

$$P = \frac{TP}{TP + FP} \qquad R = \frac{TP}{TP + FN} \qquad (6)$$

where TP is the bounding box that is correct and IOU is greater than the threshold; FP is the bounding box that is wrong and IOU is less than the threshold; FN is the bounding box, which is correct but not detected by the model. IOU is the intersection ratio between the ground truth and the bounding box predicted by the model.

Put the training sets into the model for training. After experimental analysis, the initial learning rate is set to 0,01, the batch-size is set to 6, the weight decay regular term is set to 0,0 005, the momentum value is set to 0,9, the IOU is set to 0,5, and the training epoch is set to 400.

Put the testing sets into the trained model for detection, and get the PR curve referring to the formula (6), as shown in Figure 6. The mAP is the proportion of the area enclosed by the curves and the coordinate axes in the figure. The larger the area enclosed, the better the

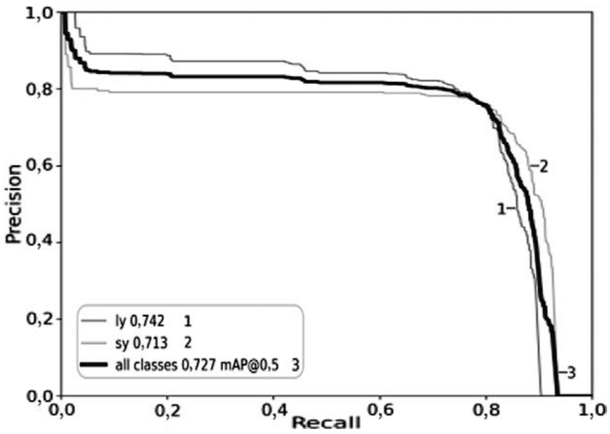


Figure 6 PR curve

model is. The horizontal coordinate is the recall rate of a certain category of the defect, the vertical coordinate is the precision rate of a certain category of the defect.

As can be seen from the Figure 6, the detection effect of marking defect is better than that of watermark defect when the IOU threshold is small. When the IOU threshold is large, the detection effect of watermark defect is better than that of marking defect. At the IOU threshold value of 0,5, the overall detection effect reaches 72,7 %.

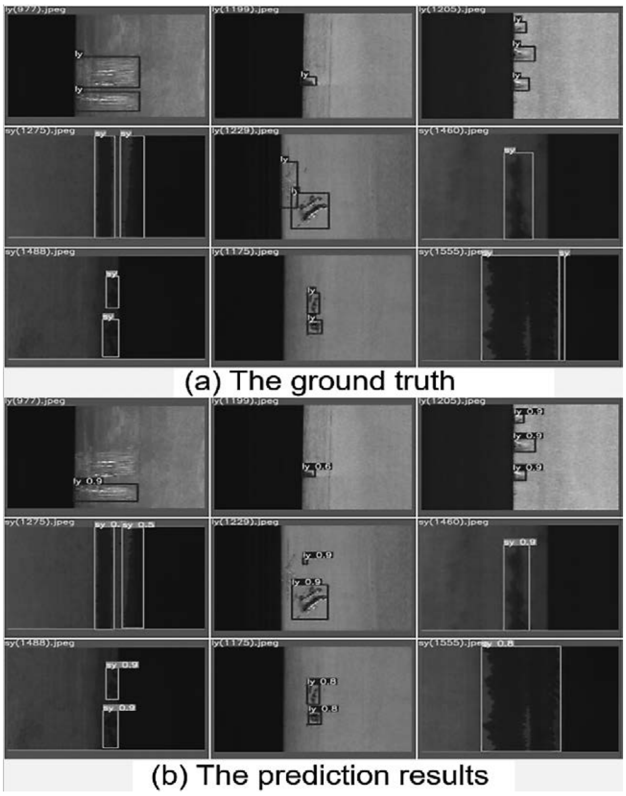


Figure 7 Defect prediction results

Some defect prediction results are shown in Figure 7. The overall prediction effect of this model is very good, and the defects can be classified and located more accurately. Finding the location of defects and clearing classification of defects in the the first time, lay a foundation for realizing the automatic production of cold rolling production line.

In the prediction stage, the detection speed of this model reaches 62 fps. Compared with the other prediction models of the two-stage methods in [3-5], the detection speed is greatly improved by 40 frames per second.

To sum up, this model has good results in mAP and fps, especially in speed.

CONCLUSION

In order to meet the production demand of online detection of steel plate surface defects, a steel plate surface defect detection method based on YOLOv3 is proposed in this paper, aiming at the problems of slow detection speed and inaccurate classification and positioning of the original method. The detection speed is up to 62 fps, and the mAP is up to 73 %. This method can not only determine whether there are defects on the surface of the steel plate, but also classify the defects. It can also accurately locate the defects, and greatly improve the detection speed. The position and category of defects in the steel plate can be monitored in time through the operation window, which is suitable for industrial production requirements, and is of great significance to realize the automatic optimization of the cold rolling production line.

Acknowledgemets

The work was financially supported by Liaoning Key Laboratory of the Internet of Things Application Technology on Intelligent Construction

REFERENCES

[1] A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks. Neural Information Processing Systems (NIPS), New York, 2012, 1097-1105

[2] Y. Li, G. Y. Li, M. M. Jiang, An End-to-End Steel Strip Surface Defects Recognition System Based on Convolutional Neural Networks Steel Research International, 88(2017)2,176-187

[3] Y. S. Weng, J. Q. Xiao, Y. Xia, Strip Surface Defect Detection Based on Improved Mask R-CNN Algorithm. Computer Engineering and Applications 57(2021)19, 235-242

[4] H. Yu, K. C. Song, Q. G. Meng, et al, An End-to-End Steel Surface Defect Detection Approach Via Fusing Multiple Hierarchical Features. IEEE Transactions on Instrumentation and Measurement 69(2019)4,1493-1504

[5] Y. Wu, Y. Z. Yang, X. L. Su, et al, Surface Defect Detection Method of Steel Plate Based on Faster R-CNN. Journal of DongHua University (Natural Science) 47(2021)3,84-89

[6] J. Redmon, A. Farhadi, YOLOv3: An Incremental Improvement. IEEE Conference on Computer Vision and Pattern Recognition, 2018

[7] K. M. He, X. Y. Zhang, S. Q. Ren, et al, Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016,770-778

Note: The responsible translators for English language is Lihua Cai – University of Science and Technology Liaoning, China