Z. H. WEI, Y. J. ZHANG, X. J. WANG, J. T. ZHOU, F. Q. DOU, Y. H. XIA

ISSN 0543-5846
METABK 63(1) 28-30 (2024)
UDC – UDK 669.14:621.771-621.41:621.9.015.004.64:620.197.6=111

A YOLOV8-BASED APPROACH FOR STEEL PLATE SURFACE DEFECT DETECTION

Received – Primljeno: 2023-04-04
Accepted – Prihvaćeno: 2023-07-20
Original Scientific Paper – Izvorni znanstveni rad

Hot-rolled steel strips are a commonly used product in both production and daily life. However, the manufacturing process inevitably leads to the occurrence of surface defects. To solve this problem, Our method uses YOLOV8 and squeeze-and-excitation (SE) attention mechanism to detect surface defects in hot-rolled steel strips. Our method balances accuracy and real-time performance, while detecting four common surface defects. The method has an average accuracy of 90,9 % and a maximum accuracy of 98,5 % for detecting a single category of surface defects. Experimental results confirm good performance of our proposed method in classifying and localizing surface defects in hot-rolled steel strips, and has the potential for broad application and promotion.

Keywords: steel strip, hot-rolled, surface defect, object detection, YOLOV8, SE attention mechanism

INTRODUCTION

Steel strip is one of the important products of China's steel industry, of which hot-rolled strip refers to strips and plates produced by hot-rolled method, generally with a thickness of 1,2~8mm. Hot rolled strip can be used directly as hot-rolled steel plate, which is widely used in automobile, motor, chemical, shipbuilding and other industrial sectors. However, due to the characteristics of high temperature and fast speed in the production process of hot-rolled steel strip, some defects on the surface of hot-rolled strip steel are often inevitable, and if not found in time, it will affect the quality of the product and even cause serious harm to the user of steel. Therefore, in modern industrial production, it is necessary to detect defects on the surface of hot-rolled steel strip.

Traditional manual detection is too subjective, and it is difficult to ensure the real-time detection. In recent years, most manufacturers have begun to use machine vision and image processing technology, and the current steel plate surface defect detection technology based on machine vision and convolutional neural network (CNN) has achieved good results. However, traditional machine vision has defects such as poor adaptability, therefore, we propose a single-stage defect detection method on the surface of steel plate based on Yolov8, which introduces new functions and improvements on the basis of retaining the basic framework, which can classify defects and determine the location of defects, which greatly improves the detection speed.

Z. H. Wei, Y. J. Zhang, X. J. Wang, J. T. Zhou, F. Q. Dou, Y. H. Xia, School of Computer Science and Software Engineering, University of Science and Technology Liaoning, China. Corresponding author: Y. J. Zhang (1997zyj@163.com)

RALATED WORK

Initially, surface defect detection and classification methods for hot-rolled steel strips relied on manually generalizing image features and detecting defects through matching methods. In [1], a representative "noise-robust method based on complete local binary patterns" was proposed, and a dataset based on NEU-DET was constructed. However, manual generalization and matching of defect images require domain experts to analyze images in specific environments for more accurate features.

In [2], a defect detection method was proposed using random forests and support vector machines. In [3], a convolutional neural network model was built using MatConvNet and GPU acceleration. In [4], a defect detection algorithm was proposed using the Swin Transformer and a multi-threshold structure.

To meet the high real-time performance requirements of hot-rolled steel strip production, this paper proposes a YOLOv8-based classification and localization method that balances real-time performance and accuracy, as the accuracy of detection greatly affects production quality. The production speed of hot-rolled steel strips can reach 400 m/min, making real-time performance crucial.

METHODOLOGY

"YOLOv8 is a state-of-the-art (SOTA) model based on the design principles of YOLOV5 and YOLOV7 ELAN as shown in Figure 1. Compared to YOLOV5, YOLOV8 introduces new features and improvements while retaining its basic framework, resulting in improved performance and flexibility. This model includes

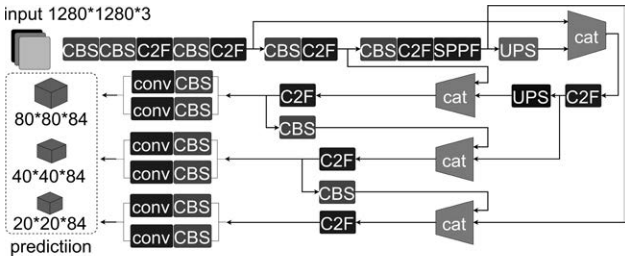

Figure 1 The network structure of YOLOV8

a new backbone network architecture, an Anchor-Free detection head, and a new loss function. Similar to YOLOV5, YOLOV8 offers models of different sizes ranging from N/S/M/L/X scales, which are adjusted based on scaling coefficients.

The backbone network and Neck sections were designed based on the YOLOV7 ELAN design philosophy. The kernel size of the first convolution layer has been reduced from 6 x 6 to 3 x 3, the C3 structure of YOLOV5 has been replaced with a more gradient-rich C2f structure, and two convolution connection layers have been removed from the Neck module. The channel numbers have been adjusted for different scale models, based on carefully designed parameters for each model to significantly improve model performance.

Compared to YOLOV5, YOLOV8 has undergone major changes in the Head section, using the popular decoupled head structure to separate the classification and detection heads, and changing the detection head from Anchor-Based to Anchor-Free. TaskAlignedAssigner positive sample allocation strategy is used in the Loss calculation, along with the introduction of the Distribution Focal Loss. In the YOLOV8 algorithm, the TaskAlignedAssigner of TOOD is directly referenced, which selects positive samples based on a score weighted by classification and regression scores."

As shown in Figure 2, the SE (squeeze-and-excitation) attention mechanism is a channel-wise attention mechanism that recalibrates channel-wise feature responses by explicitly modeling interdependencies between channels. It is proposed as a new building block that focuses on channel relationships. The SE block takes the feature map as input and first performs global average pooling based on the width and height of the feature map, which reduces spatial features to 1 x 1. Then, it establishes channel connections through two fully connected layers and a nonlinear activation function. After obtaining the normalized weights using the sigmoid activation function, the original feature map is multiplied by the weights in a channel-wise manner, which completes the channel attention process. In the first fully connected layer, the dimensionality of the feature values is reduced, which greatly reduces computation and parameters. Then, the nonlinear activation function is applied, and a fully connected layer is used to restore the original channel number, which establishes correlations between channels.

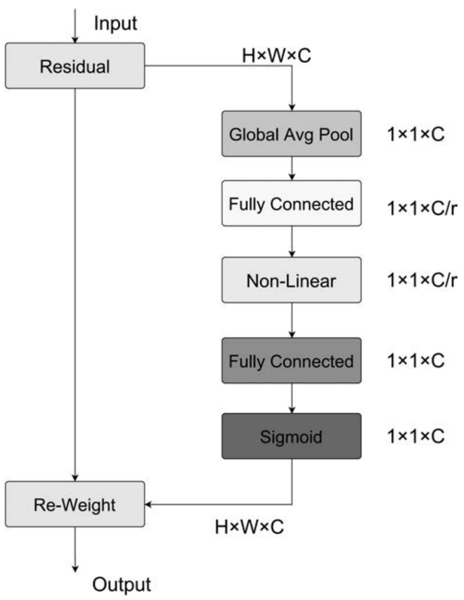

Figure 2 Structure of SE attention.

METHOD OF IMPLEMENTATION

Before training the YOLOv8 model, we select four categories in the NEU-DET dataset for preprocessing, including data cleaning and image augmentation, to expand the dataset and enhance the model's generalization performance. The dataset was then divided into a training set and a test set in an 8:2 ratio. The model's trainable parameters were randomly initialized, and the training set was batched and fed into the model for training.The model was adjusted through backpropagation using a combination of BEC Loss and CIoU Loss to calculate the error between the output results and labels. After each training round, the model's trainable parameters were fixed, and the test set was used to evaluate the model's effectiveness. This process was repeated until the model's detection performance no longer improved.

EXPERIMENT AND ANALYSIS

We selected four categories of NEU-DET steel surface defects dataset, namely: Inclusion, Patches, Pitted_surface, Scratches. The dataset's defect locations were annotated manually and the corresponding results are illustrated in the figure as shown in Figure 3.

It trained the YOLOV8 model on the NEU-DET dataset using a stochastic gradient descent algorithm with a batch size of 8. The optimizer used was SGD. The learning rate was decayed to 0,01 after 100 iterations, and the training was conducted for a total of 1,5 hours on an RTX3060 device.

$$\rho interp\left(r_{n+1}\right) = \max_{\tilde{r}:\tilde{r}\geq r_{n+1}} \rho\left(\tilde{r}\right) \tag{1}$$

$$AP = \sum_{r=0}^{1}\left(r_{n+1} - r_n\right)\rho interp\left(r_{n+1}\right) \tag{2}$$

$$mAP = \frac{\sum_{i=1}^{k} AP_i}{k} \tag{3}$$

The mAP evaluates classification and detection models based on Precision and Recall. Precision refers

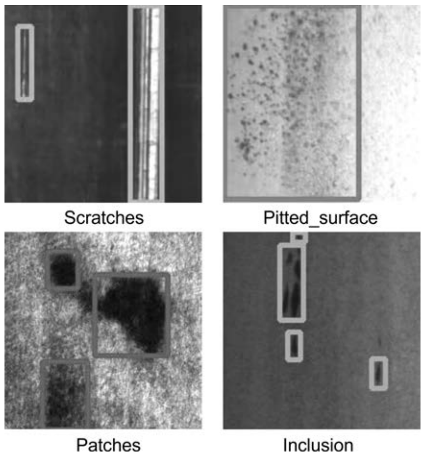

Figure 3 Annotation sample by Labelimg

to the proportion of correctly predicted positive classes among all predicted positive classes, while Recall is the proportion of correctly predicted positive classes among all positive classes in the sample. AP is calculated by averaging the Precision values corresponding to each Recall value sorted in ascending order, using formulas (1) and (2) where corresponds to the Recall value at the first interpolation of the Precision interpolation segment. Finally, mAP is obtained by averaging the AP values for each category using formula (3), where k is the number of categories.

Table 1 **Accuracy comparison between original and SE-attentive models.**

Model	Accuracy/%				
Category	Average	Inclusion	Patches	Pitted_ surface	Scratches
Original	89,5	81,0	90,8	91,7	94,6
SE	90,9	85,2	90,4	89,4	98,5

As shown in Table 1, It can be observed that the model with SE attention mechanism added demonstrates a certain level of improvement in accuracy on the dataset compared to the original model.

It used 20 % of the training set images as the test set, which were not used for training. The Precision-Recall curves and mAP values for each category and the average

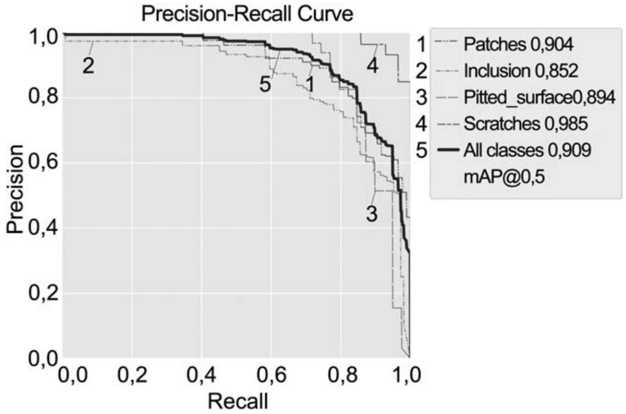

Figure 4 Precision-Recall curves for each category and mean values

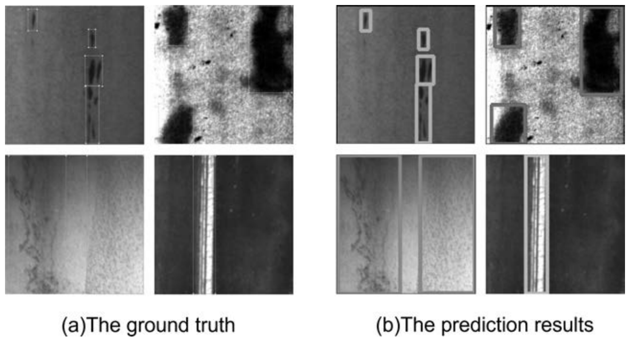

Figure 5 Defect prediction results

value of selected prediction frames with an intersection ratio greater than 0,5 are shown in Figure 4. The mAPs for the categories Inclusion, Patches, Pitted_surface, and Scratches were 85,2 %, 98,5 %, 89,4 %, and 90,4 %, respectively. The comparison between the ground truth and predicted results is illustrated in Figure 5.

The method achieves a processing speed of 59,9 images per second on an RTX3060 graphics card, with each image analysed in just 16,7ms. This is sufficient to meet the requirements for surface defect detection and localization during high-speed operation of the hot-rolled strip line.

CONCLUSION

The proposed method based on YOLOV8 achieves high accuracy and detection speed in surface defect detection and classification, and further improves the detection effect by adding the SE attention mechanism. On NEU-DET dataset, the mAP of four surface defect classification and localization tasks reaches 90,9 %, and the highest accuracy reaches 98,5 %. Its ability to process 60 images per second meets the high-speed requirements of hot rolled strip lines, and it can classify and locate surface defects in time during production.

Acknowledgements

This work was supported in part by innovation andentrepreneurship training program for collegestudents of University of science and TechnologyLiaoning 2023.

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. 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

[3] J. F. Xing. Hot rolled strip surface defect recognition and system development based on convolutional neural network.2017. MA thesis, Northeastern University.

[4] B. Yu, X. K. Zhang, and W. Wang. “Surface Defect Detection of Hot-Rolled Strip Based on STMR-CNN.” Computer Systems & Applications 31.10 (2022), 122-133.

Note: The responsible translators for English language is J. Wang – University of Science and Technology Liaoning, China