WEP
Warwick
Evans
Publishing

International Journal of Mechanical and Electrical Engineering
ISSN: 3005-9615 (Print), ISSN: 3005-7132 (Online) | Volume 3, Number 3, Year 2024
DOI: https://doi.org/10.62051/ijmee.v3n3.01
Journal homepage: https://wepub.org/index.php/IJMEE/index

Steel Surface Defect Detection Algorithm based on Improved YOLOv8

Shuai Yang

College of Technology and Education, Tianjin University, Tianjin, China

ABSTRACT

Aiming at the problem of steel surface defects, a defect detection algorithm based on YOLOv8 is constructed. Firstly, SimAM is added to the head to improve the expression ability of the features to enhance the detection ability of the model for tiny defects or small targets. Then the CIoU loss function is replaced with Wise-IoU to enhance the detection accuracy of the model. The experimental results show that the constructed improved models P, $mAP_{0.5}$ and $mAP_{0.5:0.95}$ reach 74.6%, 75.6% and 43.4%, respectively, which are improved by 5.9%, 0.5% and 0.6%, respectively, compared with the original YOLOv8n model. The detection accuracy was effectively improved.

KEYWORDS

YOLOv8; Steel Surface Defects; SimAM; Wise-IoU.

1. INTRODUCTION

As an indispensable material for infrastructure construction, machinery manufacturing and other fields, the surface quality of steel is directly related to the overall performance and safety of the product. However, cracks, scratches, plaques and other defects often appear on the surface of steel in the process of steel manufacturing and transportation. These defects directly reduce the service life of steel, and at the same time may induce potential safety hazards. Therefore, it is of great significance to improve the detection capability of steel surface defects.

Traditional steel surface defect detection methods mainly rely on manual, there are problems such as slow detection speed, high labor intensity, susceptible to human factors, which is difficult to meet the production needs of modern industry [1]. With the continuous progress of computer vision and deep learning technology, target detection algorithms based on deep learning have shown great advantages in the field of steel surface defect detection [2]. Current target detection methods can be divided into single-stage target detection and two-stage target detection. For steel surface defect detection, Jigao Xu et al [3] proposed to add CBAM attention to the infrastructure of convolutional neural network (CNN) to make the extracted features more refined. Zhao et al [4] used a multi-scale fusion training network and a deformable convolutional network to replace the traditional convolutional network on the basis of Faster-RCNN, which improves the recognition of small target defects on steel surface. capability. Kou et al [5] improved the detection accuracy and speed of end-to-end defect detection based on the YOLOv3 model. Wang et al [6] proposed a real-time steel surface defect detection network based on YOLO-v5, which introduces a multiscale exploration module to efficiently identify the multiscale information of surface defects and integrates a spatial attention module to enhance the attention to the defect information.

Although the above methods have achieved good results for steel surface defect detection, there are still some shortcomings. Since the dataset is a grayscale image with a complex background, it is difficult to distinguish the target from the background, resulting in recognition difficulties, and the small size of targets such as cracks in the defects makes such defects extremely difficult to detect. YOLO algorithm, as a real-time and fast detection target detection framework, has achieved remarkable results in the field of computer vision. In this paper, we will propose an improved algorithm based on the YOLOv8n model to increase the detection accuracy of steel surface defects.

2. YOLOV8 IMPROVED ALGORITHM

YOLOv8 is an updated and improved version of YOLOv5 by ultralytics. The official YOLOv8 has five models: YOLOv8n, YOLOv8s, YOLOv8l, YOLOv8m, and YOLOv8x. As the width and depth of the network deepens, the more complex the model structure is, the higher the detection accuracy will be. YOLOv8n model has the smallest width and depth, the fastest detection speed, in order to ensure that the model operation is controllable, so this paper chooses YOLOv8n as the basic network model.

Compared with YOLOv5, YOLOv8 replaces the C3 module with the C2f module in the backbone part, and also adjusts the number of channels to realize the lightweight of the model. The Head part separates the classification and detection heads, and uses the decoupling head to replace the detection head that performs classification and regression. The Loss side uses positive and negative matching to replace IOU matching. As a result, the YOLOv8 network model is more lightweight, with higher detection speed and detection accuracy.

In this paper, the following improvements are made to the original YOLOv8 model: firstly, SimAM (Simple Attention Mechanism) is added to the header to improve the expression ability of the features, in order to enhance the model's ability to detect tiny defects or small targets. Then the CIoU loss function is replaced with Wise-IoU to enhance the detection accuracy of the model. The improved model of YOLOv8 is shown in Figure. 1.

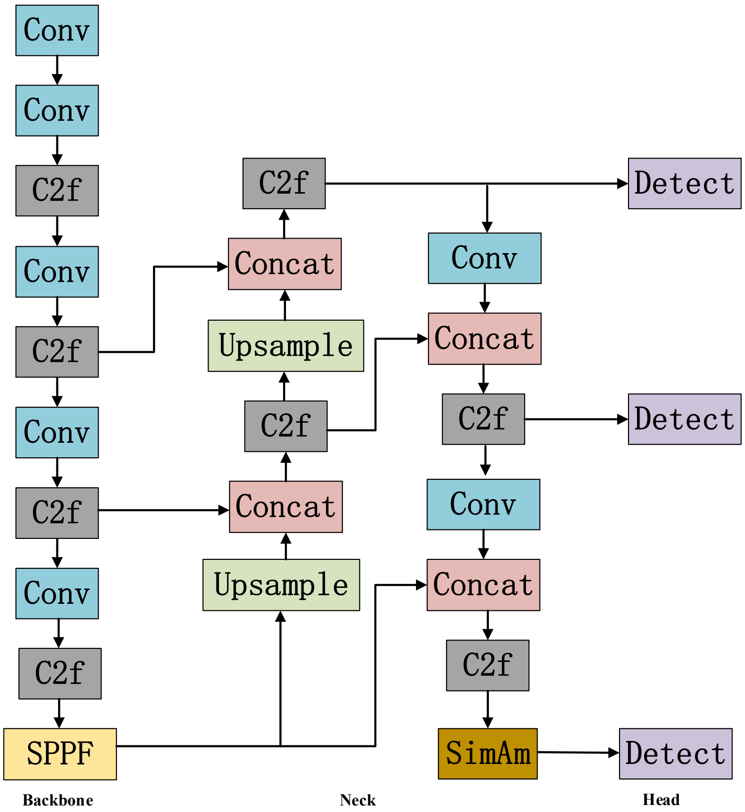


Figure 1. Improved YOLOv8 structure diagram

2.1. SimAM Module

In steel surface defect detection, fewer effective key features are extracted from small target defects such as cracks, which makes it difficult to detect. To solve this problem, this paper introduces SimAM attention mechanism in YOLOv8 header. SimAM is a lightweight, parameter-free attention mechanism, and its network structure is similar to Transformer. The structure is shown in Figure. 2. The core idea of SimAM is based on the local self-similarity of images. In an image, neighboring pixels usually have strong similarity with each other, while the similarity between distant pixels is weak. SimAM takes advantage of this property and generates attention weights by calculating the similarity between each pixel in the feature map and its neighboring pixels.

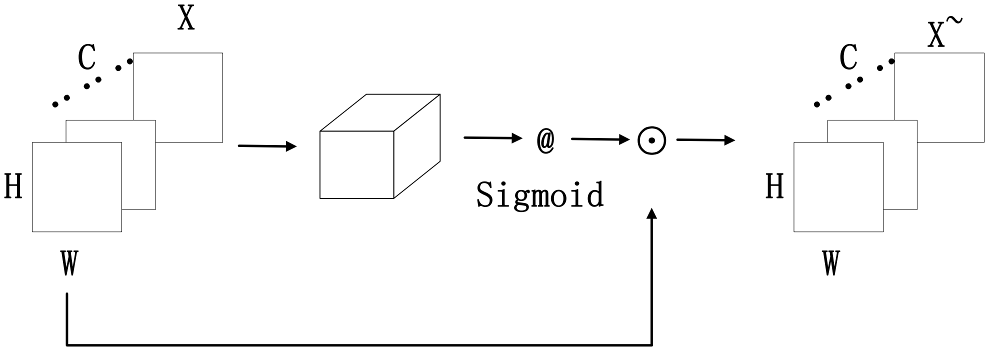


Figure 2. SimAM structure diagram

SimAM assigns higher weights to neurons with null domain inhibition to obtain neurons with strong recognition ability. The energy function e_t is defined as shown in equation (1):

$$e_t = \frac{1}{M-1} - \sum_{i=1}^{M-1}[-1-(w_t x_j + b_t)]^2 + [1-(w_t t + b_t)]^2 + \lambda w_t^2 \quad (1)$$

Where: m is the number of all neurons on a single channel; i is the index on the spatial dimension; x_i is a neuron other than the current neuron; t is the current neuron of the input feature in a single channel; w_t and b_t are the weight and bias values under the linear transformations of t and x_i, respectively; and λ is a constant, which is usually taken as IE-4.

2.2. Wise-IoU Loss Function

The YOLOv8 model uses the CIoU [7] loss function for bounding box regression and takes into account the aspect ratio of the predicted and true boxes in calculating the loss value. However, the introduction of the aspect ratio in CIoU is problematic, and either too high or too low regression quality in the regression sample may have an adverse effect on the loss function. Therefore, in this paper, the CIoU loss function is replaced with Wise-IoU to improve the detection of steel surface defects.

The dataset contains low-quality samples such as cracks and rolled oxides, so geometric factors such as distance and aspect ratio can exacerbate the penalty for low-quality samples, thus reducing the generalization performance of the model. When the prediction frame overlaps with the truth frame, a good loss function will weaken the penalty of geometric factors, and less training intervention will lead to better generalization ability of the model [8]. On this basis, the distance-attention formula is constructed to obtain the WIoU with a two-layer attention mechanism:

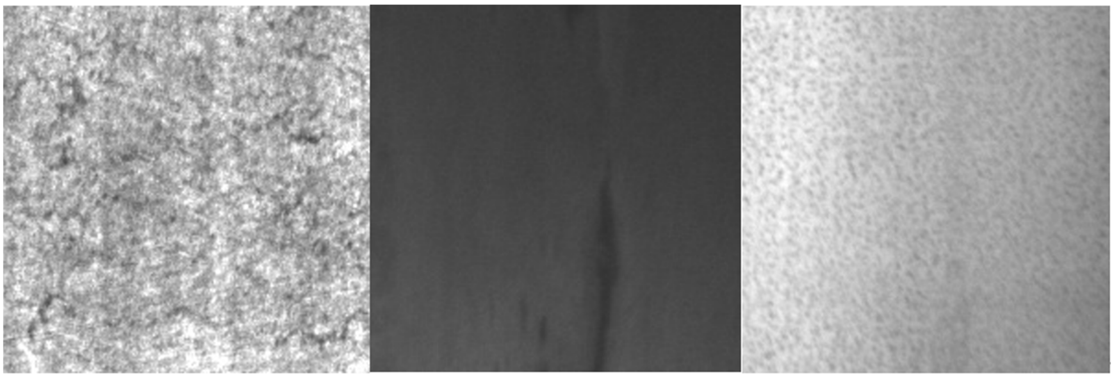
$$\begin{aligned} L_{WIoU} &= R_{IoU} L_{IoU} \\ R_{WIoU} &= \exp\left(\frac{(x-x_{gt})^2 + (y-y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \end{aligned} \quad (2)$$

Where W_g, H_g are the dimensions of the minimum detection frame.

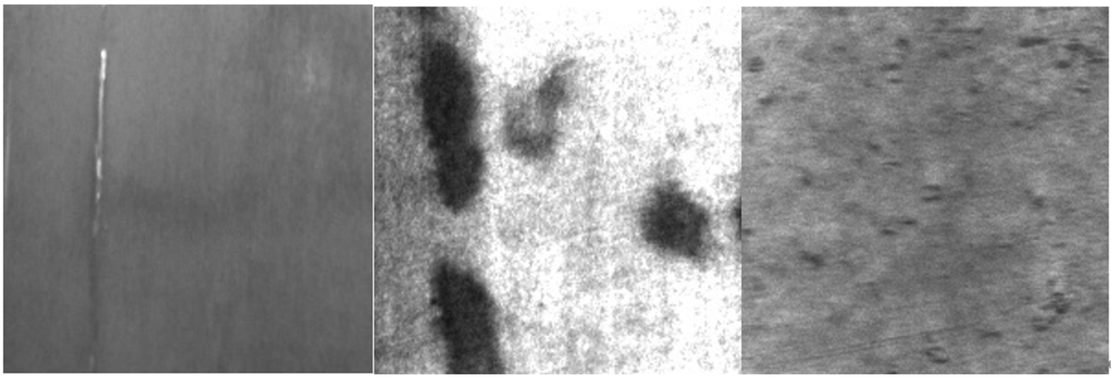
3. EXPERIMENT

3.1. Dataset Construction and Experimental Environment

This experimental dataset is derived from the online open-source dataset, including six types of defects, namely cracks, inclusions, pitting surfaces, scratches, plaques, and rolled oxides, with a total of 1400 pictures. The images are divided into a training set and a validation set according to the ratio of 7:3, with 980 pictures in the training set and 420 pictures in the validation set. An example of the dataset is shown in Figure. 3. The input image size is 640×640, batch size is set to 16, and the number of iterations is 200 rounds.



(a) Cracks (b) Inclusions (c) Pitting surfaces



(d) Scratches (e) Plaques (f) Rolling oxides

Figure 3. Categorical data sets

To ensure the fairness of the experiments, all experiments were conducted on the Ubuntu operating platform. the CPU was a 12 vCPU Intel(R) Xeon(R) Platinum 8255C CPU @ 2.50 GHz; the GPU was an Nvidia GeForce RTX 3080 10 GB; the RAM was 40 GB; and the CUDA version was 11.3; The experimental environment is Pytorch 1.11.0, Anaconda_python3.8.

3.2. Evaluation Indicators

In this paper, Precision (P) and mean average precision mean mAP (IoU thresholds are taken as 0.5 and 0.5 to 0.95) are used to evaluate the model improvement performance.

P is the proportion of positive samples of detected targets, as shown in Equation (3):

$$P = \frac{TP}{TP + FP} \tag{3}$$

The mAP is the mean of the average accuracy (AP) of each category, as shown in equations (4) and (5):

$$AP = \frac{TP + TN}{N} \tag{4}$$

$$\text{mAP}=\frac{\sum_{i=1}^{N} AP_i}{N} \tag{5}$$

Where, TP denotes the number of correctly classified in all positive cases; FP denotes the number of incorrectly classified in all positive cases; FN denotes the number of incorrectly classified in all negative cases. n is the total number of detected targets (N=TP+TN+FP+FN); and AP_i denotes the actual accuracy rate of defects in the ith category.

3.3. Comparison Experiment

In order to further verify the model performance, the improved model is compared with the original model of YOLOv8 for experiments, and the results are shown in Figure 1. It can be seen that the improved models P, $mAP_{0.5}$ and $mAP_{0.5:0.95}$ constructed in this paper reached 74.6%, 75.6% and 43.4%, respectively, which were improved by 5.9%, 0.5% and 0.6%, respectively, compared with the original YOLOv8n model. A comparison of the detection results is shown in Figure 4.

Table 1. Comparison experimen

Model	P(%)	$mAP_{0.5}$(%)	$mAP_{0.5:0.95}$(%)
YOLOv8n	68.7	75.1	42.8
Model of this paper	74.6	75.6	43.4

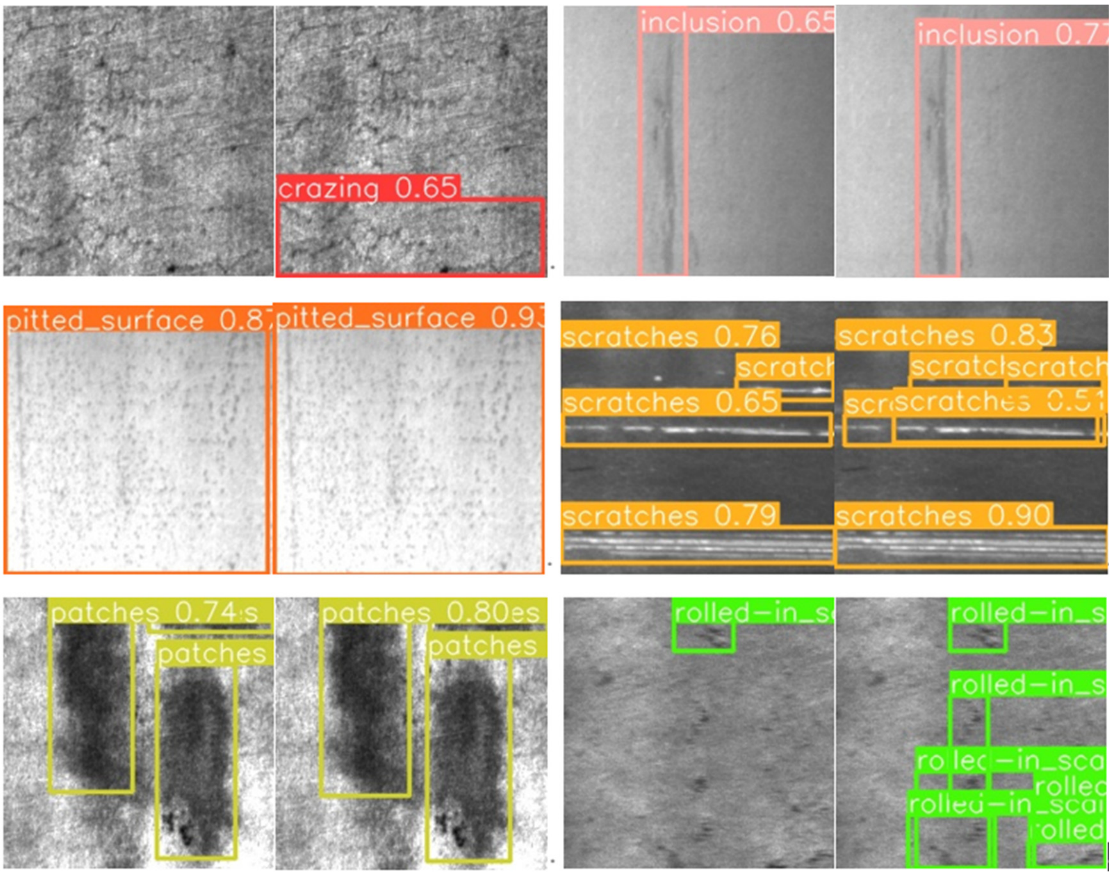


Figure 4. Comparison effect before and after model improvement

Among them, the accuracy improvement is significant, and Table 2 shows the comparison of the accuracy of various types of defects. It can be seen that the original detection accuracy is low due to the cracks and rolling oxide defects and the background color is similar, resulting in difficult to identify, so the accuracy of the original detection is low, the accuracy of the improved model constructed in this paper has been improved by 15.4% and 4.6%, respectively, which is a significant effect. The detection accuracies of other defects are also improved.

Table 2. Comparison of accuracy of various types of defects

Defect type	Crazing	Inclusion	Pitting_surface	Scratches	Patches	Rolled-in-scale
P(YOLOv8n)	54.3	71.1	78.6	69.5	81.8	56.9
P (this paper)	69.7	73.4	84.5	73.1	84.8	64.2

4. CONCLUDING REMARKS

In this paper, an improved steel surface defect detection method is proposed in the framework of YOLOv8n. The improved models P, $mAP_{0.5}$ and $mAP_{0.5:0.95}$ constructed in this paper reach 74.6%, 75.6% and 43.4%, respectively, which are improved by 5.9%, 0.5% and 0.6%, respectively, compared with the original YOLOv8n model. Although the improvement in detection accuracy is good, the overall fine-tuning accuracy is low due to the large number of defect types and the difficulty of detecting certain defect types, so the next step will be to study how to better detect cracks and other defects with smaller targets.

REFERENCES

[1] Changlu Xu, Linsheng Li, Jiwei Li. “Defect detection method for lithium battery pole piece based on the fusion of Gamma correction and LOG algorithm”, Mechanical Design and Research, Vol. 36, p. 105-109, 2020.

[2] W. W. Li, Q. Li, C. Gao. “An unmanned vehicle target detection algorithm based on multi-scale feature fusion”, Journal of Beijing University of Information Science and Technology (Natural Science Edition), Vol. 36, p. 19-23, 2021.

[3] Jigao Xu, Jiangyi Chen, Dongchen Qin, et al. “Research on steel surface defect detection based on CNN and CBAM”, Mechanical Design and Manufacturing, p. 1-5, 2024. https://doi.org/10.19356/j.cnki.1001-3997.20240725.016.

[4] Zhao W, Chen F, Huang H, et al. “A new steel defect detection algorithm based on deep learning”, Computational Intelligence and Neuroscience, p. 1-13, 2021.

[5] Kou X, Liu S, Cheng K, et al. “Development of a YOLO-V3-based model for detecting defects on steel strip surface”, Measurement, Vol. 182, p.109454, 2021. https://doi.org/10.1016/j.measurement.2021.109454.

[6] Wang L, Liu X, Ma J, et al. “Real-Time Steel Surface Defect Detection with Improved Multi-Scale YOLO-v5”, Processes, Vol. 11, p. 1357, 2023.

[7] ZHENG Z,WANG P,LIU W, et al. “Distance-IoU loss: Faster and better learning for bounding box regression”, arXiv, 2019.DOI:10.1609/aaai. v34i07.6999.

[8] Li Y T, Fan Q S, Huang H S, et al. “A modified YOLOv8 detection network for UAV aerial image recognition”, Drones, Vol. 7, p. 304, 2023.