

# Research on target defect detection algorithm based on improved YOLO-V7

Xingchen Zhang

Zhengzhou Univeristy, Henan 450001, China

**Abstract.** The goal of this study is to increase target identification accuracy and defect detection performance using the enhanced YOLO-V7 target defect detection algorithm. You Only Look Once version 7 is referred to as YOLO-V7 and is a well-liked real-time target identification technique. To make high-quality goods, however, it is essential to find flaws, and the traditional YOLO-V7 has certain restrictions when it comes to addressing specific flaws. To get around these restrictions, we implemented a number of changes to YOLO-V7. The study's enhanced YOLO-V7 target defect detection algorithm may find use in the areas of industrial automation, quality assurance, and safety monitoring.

**Keywords:** YOLO-V7; Target defect detection; Slou; Iron and steel surfaces; Quality control.

## 1. Introduction

Target defect detection is of great significance in the fields of industrial automation, quality control and safety monitoring. With the rapid development of computer vision and deep learning, target detection algorithms have made remarkable progress. Among them, You Only Look Once version 7 (YOLO-V7), as an efficient real-time target detection algorithm, has attracted wide attention and applications.

However, the traditional YOLO-V7 algorithm has some challenges and limitations in target defect detection. For example, when dealing with complex defect scenarios, the traditional YOLO-V7 may suffer from inaccurate localization, missed detection, and false detection, which affects the accuracy and stability of detection[1]. Therefore, it is necessary to improve YOLO-V7 to enhance its performance in target defect detection tasks.

Steel surface defect detection is one of the important tasks in industrial production and quality control. The presence of steel surface defects may affect the performance and quality of steel, so early detection and accurate testing of these defects is essential to ensure product quality[2]. The following are some common methods for detecting steel surface defects:

**Visual inspection methods:** Visual inspection is a traditional and widely used method. The steel surface is photographed or scanned by a camera or optical sensor, and then the image is processed and analyzed using computer vision algorithms[3]. Common algorithms include edge detection, morphological processing, image segmentation and feature extraction. Visual inspection methods work better for some simple defect detection, but more complex algorithms may be needed for complex defects.

**Magnetic Particle Inspection Method:** Magnetic particle inspection is a commonly used method for detecting cracks on the surface of steel. It utilizes the magnetic properties of ferromagnetic materials, and after applying a magnetic field[4], magnetic powder will gather at the surface cracks, forming visible magnetic powder accumulation lines. This method is more effective for detecting surface cracks parallel to the direction of the magnetic field.

**Ultrasonic Inspection Methods:** Ultrasonic inspection is a method commonly used to detect internal and surface defects in steel[5]. This is accomplished by sending ultrasonic waves to the surface or interior of the steel and then receiving the signals that are reflected back. If there is a defect, the ultrasonic wave will be scattered or reflected by the defect, thus forming an echo signal, and by analyzing the characteristics of the echo signal, we can determine whether there is a defect or not[6].

**Eddy current inspection method:** Eddy current inspection is a method used to detect surface defects in conductive materials, and is suitable for detecting thin-layer surface defects in steel. It utilizes the

principle of eddy current induction to generate eddy currents by applying alternating currents on the steel surface. When the eddy current encounters a surface defect, it causes a change in the current, and the defect can be detected by detecting this change.

**Infrared Thermal Image Inspection Methods:** Infrared Thermal Image Inspection utilizes an infrared camera to detect the heat distribution on the steel surface, as defects usually cause local temperature changes. Through infrared image processing and analysis, thermal anomalies on the surface can be detected to find defects[7].

The combined application of the above methods can improve the accuracy and efficiency of steel surface defect detection. Different methods are suitable for different types of defects, and usually a combination of methods is used in practice for comprehensive detection. At the same time, with the continuous development of science and technology, artificial intelligence and deep learning and other new technologies are gradually applied to the field of steel surface defect detection, providing more intelligent and efficient solutions for defect detection.

## 2. Related Work

Prior to the study of improved YOLO-V7 based algorithms for the task of target defect detection, many related works have been explored in the field of target detection and defect detection. The following are some of the works related to this study:

In the field of target detection, YOLO-V7 is an evolutionary version of You Only Look Once version 7. The YOLO family of algorithms is a class of end-to-end real-time target detection algorithms that predicts the bounding box and class probabilities of a target directly in a single forward pass by treating the target detection task as a regression problem. Prior to YOLO-V7, versions such as YOLO-V6 have achieved significant performance improvements. However, these traditional YOLO algorithms have some limitations when dealing with target defect detection and need further improvement.

A variety of algorithms have been proposed by many researchers for the task of target defect detection. Some of these traditional methods use hand-designed features and classifiers, such as SVM (Support Vector Machine), Adaboost (Adaptive Augmentation Algorithm), etc., to achieve defect detection. However, these methods may not perform well in complex defect scenarios[8].

In recent years, the rise of deep learning techniques has brought new opportunities for target defect detection. Some research works have used Convolutional Neural Network (CNN)-based methods such as Faster R-CNN, SSD (Single Shot Multibox Detector), and Mask R-CNN for target defect detection[9]. These methods improve the performance of traditional algorithms to some extent, but there are still some challenges.

In deep learning, data augmentation is a commonly used technique to increase the diversity and quantity of data by randomly transforming and expanding the training data. Data augmentation can improve the generalization ability of the model and is especially important in target defect detection for the case of insufficient defect samples.

Based on the research status and problems of the above related work, this study will work on improving the performance of target defect detection algorithms based on the improved YOLO-V7 algorithm, combined with a more powerful feature extraction network, a new loss function design, and data enhancement techniques. By evaluating the algorithm on real datasets, we hope to prove the superiority of the improved algorithm in the task of target defect detection and provide a more reliable solution for fields such as industrial production and quality control.

## 3. Modeling

### 3.1 Model YOLO-V7

The YOLO-V7 is a reference to the YOLO series, and it is faster and more accurate than any other object detectors between 5 and 160 frames per second. The YOLO-V7 strikes the ideal mix between

speed and precision, earning it recognition within the industry. Backbone, encoder, and decoder are the three components that make up its overall detection pipeline. The three primary components of the YOLO-V7 structure are input, the backbone feature extraction network, the strengthened feature extraction network, and predictions.

### 3.2 YOLO-V7 Model-Based Improved Network Structure

This approach suggests a novel steel strip surface defect detection technique that uses YOLO-V7 as the baseline. As a total, YOLO-V7 reduces the input picture to 640\*640, feeds it into the backbone network, produces three layers of feature maps of various sizes via the head network, and finally feeds RepConv with the prediction outcome. ELAN, MP structures, and Silu activation function are the major components of the YOLO-V7 backbone. By regulating gradient routes and deeper networks, the ELAN structure is able to learn and converge effectively. ELAN-W is comparable. Downsampling use the MP structure. The ECA attention mechanism was introduced to the backbone's base. A portion of the SE attention process is followed by this module. The one-dimensional convolution kernel's size is adaptively chosen, and the dimension is maintained throughout local cross-channel interaction, which reduces network complexity and boosts model performance. This is the key advancement over the SE attention technique. In ECA, the original feature image is first input, after which all of the image's channels are globally averaged and pooled. Next, channel weights are produced using a quick one-dimensional convolution with a size of Q, and the corresponding probabilities of the various channels are calculated and then compared to the original image[10]. As the input to the next layer, the input characteristics are multiplied collectively. According to formula (1) (2), this approach uses function adaptation to get the Q value, whose value is proportional to the channel dimension C.

$$C = \phi(Q) = 2^{(\lambda Q - b)} \quad (1)$$

$$Q = \psi(C) = \left\lfloor \frac{\log_2(C)}{\lambda} + \frac{b}{\lambda} \right\rfloor \text{ odd} \quad (2)$$

when  $\lambda = 2$ ,  $b=1$ , Q selects the closest odd integer. ECA has an adaptable and light structure that enables adaptive one-dimensional convolution kernel selection, direct cross-channel communication without dimensionality reduction, improvement of useful semantic information in feature maps, and the effective extraction of steel surface defect features. ECA boosts YOLO-V7's effectiveness and is appropriate for the datasets used in this paper[11].

The original BiFPN structure, which is often employed to enhance feature fusion, provides various weights to input characteristics based on their relative significance. However, we discovered that adding weighted BiFPN to YOLO-V7 does not provide the best results. We postulate that this is because weighing input feature layers and adding attention techniques are both very comparable procedures. As a result, we remove the weight component and produce a dewighted BiFPN.

### 3.3 Loss Function

The bounding box loss function, the objectness loss function, and the class loss function make up the YOLO-V7 loss function. The error of the prediction box for the coordinate positioning error is measured using the bounding box loss function. The confidence error of the prediction box is reflected in the objectness loss function. The prediction error of the prediction box for the target category is reflected in the class loss function[12]. BCEWithLogitsLoss is the YOLO-V7's objectness loss function and class loss function. CIoU loss is the objectness loss function. The bounding box regression is more stable when CIoU takes into account the distance between the ground truth box and the prediction box, the overlap rate, the box scale, and the penalty term. Equation (3) gives a definition of CIoU loss.

$$\text{Loss}_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (3)$$

where  $\rho^2(b, b^{gt})$  stands for the Euclidean distance between the centers of the prediction and ground truth boxes, and it is denoted. Where  $c$  is the minimal closure rectangle's diagonal distance,

which contains both the prediction box and the ground truth box. Equation (4) gives a definition of. Eq. (5) gives the definition of  $\alpha$ .

$$\alpha = \frac{v}{1-\text{IoU}+v} \tag{4}$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \tag{5}$$

Where  $w^{gt}$  represents the width of the ground truth box. where hgt stands for the ground truth box's height. where w stands for the prediction box's width. where h stands for the prediction box's height. The orientation between the ground truth box and the prediction box is not taken into account while analyzing bounding box loss functions, such as CIoU, which results in a sluggish convergence rate. In order to redefine the correlation for this, SIoU adds the vector angle between the prediction box and the ground truth box. Thus, the SIoU loss function was used in lieu of the CIoU loss function.

**Table 1.** AP results of different models on GC10-DET.

Types	SSD300	SSD512	Faster-RCNN	YOLO-V5	YOLO-7	Proposed Method
Pu	0.860	0.891	0.899	0.953	0.93	0.969
Wl	0.974	0.885	0.554	0.877	0.681	0.798
Cg	0.861	0.848	0.872	0.96	0.986	0.991
Ws	0.552	0.558	0.599	0.76	0.905	0.908
Os	0.612	0.662	0.653	0.627	0.883	0.885
Ss	0.689	0.650	0.579	0.687	0.891	0.91
In	0.168	0.256	0.194	0.367	0.782	0.869
Rp	0.105	0.364	0.364	0.539	0.848	0.826
Cr	0.527	0.521	0.736	0.427	0.83	0.868
Wf	1.000	0.919	0.818	0.822	0	0
mAP	0.635	0.651	0.627	0.702	0.774	0.802
FPS	29.3	27.1	23.6	-	111.1	105.2

A comparison of AP, mAP, and FPS for the defects on the GC10-DET dataset and NEU-DET dataset is shown in Tables 1 and 2. The detection approach based on YOLO-V7 is much faster and more accurate than previous models on the GC10-DET dataset. The enhanced YOLO-V7 model outperforms the others in terms of AP for the seven different kinds of flaws, including Pu, Cg, Ws, Os, Ss, In, and Cr. Although somewhat slower than the original YOLO-V7 model, the detection speed is also much quicker than other models and may be disregarded at high rates of FPS > 100. But the YOLO-V7 model does terribly on flaws with lighter color, like Wf. Comparative tests are carried out using the NEU-DET dataset to further confirm the robustness of the technique. The revised YOLO-V7 model has the best detection rates for Cr, Rs, and Sc defects on the NEU-DET dataset. It is also substantially quicker than other versions and just marginally slower than the original YOLO-V7 model in terms of detecting speed. The experimental findings show that the model described in this work is more accurate and quick at detecting steel surface defects than previous models.

**Table 2.** shows the AP outcomes for several models on NEU-DET.

Types	SSD300	SSD512	Faster-RCNN	YOLO-V5	YOLO-7	Proposed Method
Cr	0.411	0.417	0.374	0.509	0.517	0.69
In	0.796	0.763	0.794	0.855	0.861	0.853
Pa	0.839	0.863	0.853	0.948	0.932	0.907
Ps	0.839	0.851	0.851	0.829	0.811	0.843
Rs	0.621	0.581	0.545	0.617	0.64	0.727
Sc	0.836	0.856	0.882	0.874	0.874	0.896
mAP	0.714	0.724	0.711	0.772	0.773	0.819
FPS	37.6	29.0	23.8	-	58.8	55.6

## 4. Conclusion

Through this study, we successfully achieved significant performance improvement based on the improved YOLO-V7 target defect detection algorithm. Based on the traditional YOLO-V7 algorithm, we introduce a more powerful feature extraction network, adopt a new loss function design, and apply data enhancement techniques to improve the accuracy and robustness of target defect detection.

Extensive experimental evaluations on common target defect datasets show that the improved YOLO-V7 algorithm exhibits significant advantages in the target defect detection task. Compared to the traditional YOLO-V7 algorithm and other target detection algorithms, our algorithm exhibits higher accuracy and better performance in target defect detection. This indicates that our algorithm is able to locate and detect defects on steel surfaces more accurately, effectively improving the accuracy and stability of defect detection.

At the same time, our algorithm achieves these significant improvements without sacrificing real-time performance, which is critical for practical applications in industrial automation. Efficient real-time detection of steel surface defects will help detect potential quality problems in advance and reduce the generation of defective products, thereby reducing production costs and improving product quality.

However, we also realize that there are still some challenges in the research of target defect detection algorithms. For example, the detection accuracy for small-size defects still needs to be further improved, while more algorithm optimization and tuning may be required when applied to other complex scenarios. In addition, the quality and diversity of the dataset also have an impact on the performance of the algorithm, so we will continue our efforts to collect more high-quality data and improve the data enhancement strategy.

In future research, we will further optimize the improved YOLO-V7 algorithm to further improve the accuracy and robustness of target defect detection. Meanwhile, we will also explore more deep learning techniques and data enhancement methods in order to achieve better results in a wider range of application scenarios. Based on the research results of the improved YOLO-V7 algorithm, we believe that the target defect detection technology will be more widely used in the fields of industrial production and quality control, and contribute to the improvement of product quality and production efficiency.

## References

- [1] Ma, W., Yu, H., Fang, W., Guan, F., Ma, D., Guo, Y., ... & Wang, C. (2023). Crop Disease Detection against Complex Background Based on Improved Atrous Spatial Pyramid Pooling. *Electronics*, 12(1), 216.
- [2] Ji, C., Wang, B., Zhang, F., Huang, X., Hou, W., & Song, Z. Research on Insulator Defect Detection Algorithm in Different Environments Based on Multi-Head Attention Mechanism. Available at SSRN 4308318.
- [3] Chen, X., Yuan, M., Yang, Q., Yao, H., & Wang, H. (2023). Underwater-YCC: Underwater Target Detection Optimization Algorithm Based on YOLOv7. *Journal of Marine Science and Engineering*, 11(5), 995.
- [4] Liu, J., Zhao, H., Chen, Z., Wang, Q., Shen, X., & Zhang, H. (2023). A Dynamic Weights-Based Wavelet Attention Neural Network for Defect Detection. *IEEE Transactions on Neural Networks and Learning Systems*.
- [5] Shahin, M., Chen, F. F., Hosseinzadeh, A., Khodadadi Koodiani, H., Bouzary, H., & Shahin, A. (2023). Enhanced safety implementation in 5S+ 1 via object detection algorithms. *The International Journal of Advanced Manufacturing Technology*, 125(7-8), 3701-3721.
- [6] Feng, P., Qi, Y., Deng, S., & Zhu, C. (2022). Research on Human Recognition of Post Disaster Search and Rescue Robot based on YOLOv7 Fusionattention Mechanism. *International Core Journal of Engineering*, 8(11), 22-27.

- [7] Soeb, M. J. A., Jubayer, M. F., Tarin, T. A., Al Mamun, M. R., Ruhad, F. M., Parven, A., ... & Meftaul, I. M. (2023). Tea leaf disease detection and identification based on YOLOv7 (YOLO-T). *Scientific reports*, 13(1), 6078.
- [8] Song, Z., Huang, X., Ji, C., & Zhang, Y. (2023). Intelligent Identification Method of Hydrophobic Grade of Composite Insulator Based on Efficient GA-YOLO Former Network. *IEEJ Transactions on Electrical and Electronic Engineering*.
- [9] Ye, G., Qu, J., Tao, J., Dai, W., Mao, Y., & Jin, Q. (2023). Autonomous surface crack identification of concrete structures based on the YOLOv7 algorithm. *Journal of Building Engineering*, 73, 106688.
- [10] Yang, W., Luo, W., Mao, J., Fang, Y., & Bei, J. (2022, December). Substation meter detection and recognition method based on lightweight deep learning model. In *International Symposium on Artificial Intelligence and Robotics 2022* (Vol. 12508, pp. 199-207). SPIE.
- [11] Kumar, A. (2023). An XNOR-ResNet and spatial pyramid pooling-based YOLO v3-tiny algorithm for Monkeypox and similar skin disease detection. *The Imaging Science Journal*, 71(1), 50-65.
- [12] Kinasih, F. M. T. R., Machbub, C., Yulianti, L., & Rohman, A. S. (2023). Two-stage multiple object detection using CNN and correlative filter for accuracy improvement. *Heliyon*, 9(1).