**Frontiers in Computing and Intelligent Systems**
ISSN: 2832-6024 | Vol. 6, No. 2, 2023

# PCB Board Defect Detection Method based on Improved YOLOv8

**Chang Liu [1,*], Xiangyang Zhou [2], Jun Li [2], Chuantao Ran [1]**

[1] Department of Intelligent Manufacturing and Automotive, Chong Qing College of Electronic Engineering, Chongqing, 401331, China
[2] Chongqing Changan Automobile Company Limited, Chongqing, 400023, China

[*] **Corresponding author**: Chang Liu (Email: 744620708@qq.com)

**Abstract:** This study provides an improved YOLOv8-based printed circuit board (PCB) defect identification method to address the current challenges associated with PCB defect detection, including the detection of small targets, low accuracy, and other related concerns. The YOLOv8 model serves as the foundational framework, and in order to enhance detection speed, the YOLOv8s model is selected due to its reduced parameter count. However, feature extraction becomes challenging for small target defects; to address this, the CA attention mechanism is implemented, which is more attuned to target feature information and aids in feature extraction. As indicated by the experimental findings, the enhanced YOLOv8s-CA algorithm model has the following characteristics: a footprint of 5.79 MB, a mean average precision (mAP) of 90.4 percent, an increase of 6.6 percent over the initial network, and a parameter count augmentation of merely 0.007M. Consequently, this model finds utility in compact industrial inspection apparatus and possesses a wide range of potential applications.

**Keywords:** PCB Board Inspection; YOLOv8; Defect Detection; CA Attention Mechanism.

## 1. Introduction

Ever since the advent of the industry 4.0 era, electronic devices have become an integral part of our daily lives. PCB, an essential electronic component, serves as a conduit for connecting diverse electronic components and supplying equipment with physical support and circuit connections. Almost all electronic devices are PCB-required [1–2]. In industrial production, however, PCB faults may result from poor storage procedures, operating mode malfunctions, or soldering errors. Current PCB flaws can be broadly classified into six distinct categories: leakage, short circuit, open circuit, stray copper, and mouse bites [3]. To guarantee the integrity and security of electronic equipment, the detection of surface flaws on PCBs is an essential task [4–5].

Presently, deep learning algorithm-based PCB flaw detection technology development is relatively advanced. One major category comprises related algorithms, the other of which is the first-level target detection method shown by the SSD (Single Shot MultiBox Detector) [6], YOLO (You Only Look Once) series [7–9], and others. The second is Faster R-CNN (faster region-based CNN), which represents a two-stage method for target detection [10], and so on. In order to enhance PCB's ability to recognize small targets, the authors [9] implemented a small target detection layer and depth-separable convolution to provide real-time high-precision detection and model volume compression. Furthermore, the detection of prevalent PCB defect types was successfully concluded by Hu et al. [10], who enhanced Faster-RCNN and implemented ResNet50 with an image pyramid as a feature extraction network.

While employing deep learning target detection methods to identify PCB faults can enhance detection accuracy and efficiency, there are still obstacles that must be surmounted. The diversity of fault types is the initial: PCB flaws encompass a wide variety of forms [11], such as welding defects, circuit defects, and component installation defects, among others. The identification of diverse types of faults necessitates the development of suitable network architectures and loss functions, as their properties and morphologies may vary. The second is the real-time need [12]: Defect detection in the PCB production process is typically required to occur in real-time so that problems can be identified and addressed promptly. In order for deep learning models to satisfy real-time demands, their inference speeds must be quick. It is necessary to optimize and accelerate the model using technologies like quantization, compression, and others.

Given the multitude of challenges encountered in the aforementioned research regarding the detection of surface defects on PCB boards, the CA attention [15] mechanism was implemented in lieu of the SE attention mechanism in order to improve the model space recognition capability, utilizing the YOLOv8s target detection model [13-14]. Accelerate model convergence, enhance the model's detection effect on minute target flaws, and improve the accuracy of model detection. The experimental findings demonstrate that the enhanced model is capable of detecting PCB surface defects with more precision, which is advantageous for the implementation of industrial inspection systems that identify PCB board defects.

## 2. Introduction to YOLOv8-CA Model

### 2.1. CA-based Model Optimization

The following impacts may result from including the coordinate attention mechanism into the target detection model for faulty tiny targets: By assigning weights to the critical positions of flawed small targets during the detection process, the coordinate attention mechanism can reduce localization error and improve localization accuracy for small targets. Additionally, it enhances the region of interest by directing the model's focus towards the critical regions of defective small targets, thereby increasing the sensitivity of small target detection. region, enhance the sensitivity of small target detection, and then adapt to targets of different scales.

Since defective small targets are typically small in comparison to the entire image, the coordinate attention mechanism can adjust the weight of attention adaptively based on the target's scale information, allowing the model to adapt more effectively to defective small targets of different scales.

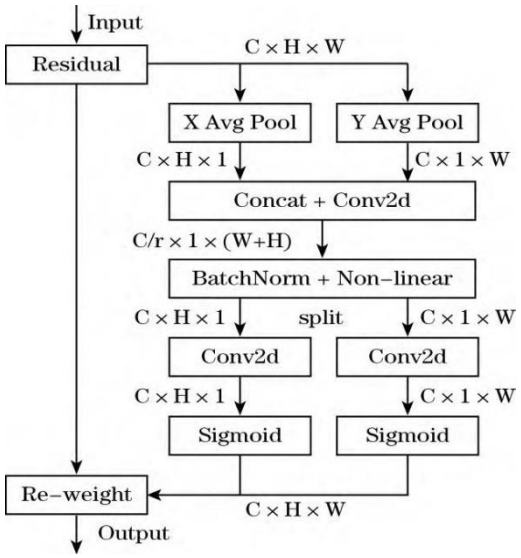


**Figure 1.** CA attention mechanism structure diagram

The structure of the CA attention mechanism is shown in Figure 1, and the CA module process consists of the following two steps:

(1) Coordinate information embedding

The use of the global pooling method by the channel attention modules (SENet, ECANet) in the global encoding of spatial information will result in the loss of location information. In order to solve the above problem and enhance the ability to capture location information, global pooling is differentiated with reference to Equation 1 and transformed into a coding operation in the form of one-dimensional features:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x_c(i,j) \tag{1}$$

Precisely, a Pooling kernel of size (H, 1) or (1, W) is employed to encode the channel along the X-axis and Y-axis for a given input. This encoding enables the expression of the output of the first c channels at height h as Equation 2.

$$z_c^h(h) = \frac{1}{W} \sum_{0 \le i < W} x_c(h,i) \tag{2}$$

Likewise, Equation 3 describes the output of the first c channel with a width of w.

$$z_c^w(w) = \frac{1}{H} \sum_{0 \le j < H} x_c(j,w) \tag{3}$$

The aforementioned transformations aggregate features in the horizontal and vertical spatial directions, respectively. The resultant feature maps are cognizant of orientation and assist the network model in identifying the specific area of interest within the input image.

(2) Coordinate Attention Generation

A second alteration is performed on the characteristics produced by coordinate information embedding; this operation is known as coordinate attention generation. Following the information embedding operation to convert the features, the concatenate operation will be executed in this section. Subsequently, the conv transforms function F1 (with a size of 1×1) will be employed to execute the subsequent operation on the features:

$$f = \delta\left(F_1\left(\left[z^h, z^w\right]\right)\right) \tag{4}$$

Equation 4, $[\cdot,\cdot]$ represents the Concatenate operation of feature information along the spatial dimension, $\delta$ represents the nonlinear activation function, and $f$ represents the intermediate feature map that encodes spatial information in both horizontal and vertical directions. $r$ is used to control the reduction rate of SE module size. Next, $f$ is divided into $f^h \in R^{C/r*H}$ and $f^w \in R^{C/r*H}$ according to the spatial dimension. Finally, two 1×1 convolution transformations $F_h$ and $F_w$ are used to transform $f^h$ and $f^w$ respectively into tensors with the same number of channels and input them to X, resulting in formulas 5 and 6:

$$g^h = \sigma\left(F_h\left(f^h\right)\right) \tag{5}$$

$$g^w = \sigma\left(F_w\left(f^w\right)\right) \tag{6}$$

Equation 5 and Equation 6, $\sigma$ represents the sigmoid activation function. In order to further reduce the weight of the model, an appropriate reduction ratio of $r$ is selected to reduce the number of channels to $f$, and then the outputs $g^h$ and $g^w$ are expanded respectively and used as Attention Weights.

To sum up, the output $Y$ of the CA module can be expressed as:

$$y_c(i,j) = x_c(i,j) \times g_c^h(i) \times g_c^w(j) \tag{7}$$

By applying horizontal and vertical attention to the input tensor via the CA module, it is possible to verify that every element in both feature representations is present inside the region of interest. By employing this mode of operation, the CA module is capable of precisely determining the positions of the effective features, thereby facilitating the precise detection of a wide range of flaws on PCB boards.

## 2.2. Introduction to the YOLOv8 Model

As one of the traditional single-stage detection methods, the YOLO algorithm has been enhanced to YOLOv8. This work presents an enhanced approach for YOLOv8s that eliminates the need for model size, computation, and parameter counts while simultaneously improving detection accuracy.

The components of the network are illustrated in Figure 2, utilizing the structure of YOLOv8 as a case study. The backbone network comprises the CBS, C2f, and SPPF modules, employing the same concept as CSP. However, the neck network of the backbone network adheres to the ELAN structure design principle of YOLOv7 and substitutes the C3 module in YOLOv5 with the C2f structure, which features a more pronounced gradient flow. This modification enhances the convolutional neural network's capability to fuse features and accelerates the rate of inference.

The YOLOv8 design continues to employ the SPPF module, which was implemented in the YOLOv5 architecture. This SPPF module functions as a spatial pyramid pooling layer, facilitating the fusion of local and global features and

augmenting the feature information. Compared to YOLOv5, the head section has undergone significant changes and has been replaced with the standard decoupled head structure, in which the classification and detecting heads are separated. YOLOv8 replaces the Anchor-Base approach with the Anchor-Free approach; for the loss function, it employs BCELoss for classification loss and DFLLoss+CIOU Loss for regression loss; and for sample matching, it swaps the Task-Aligned Assigner method for the IOU matching or unilateral proportion method that was previously utilized.

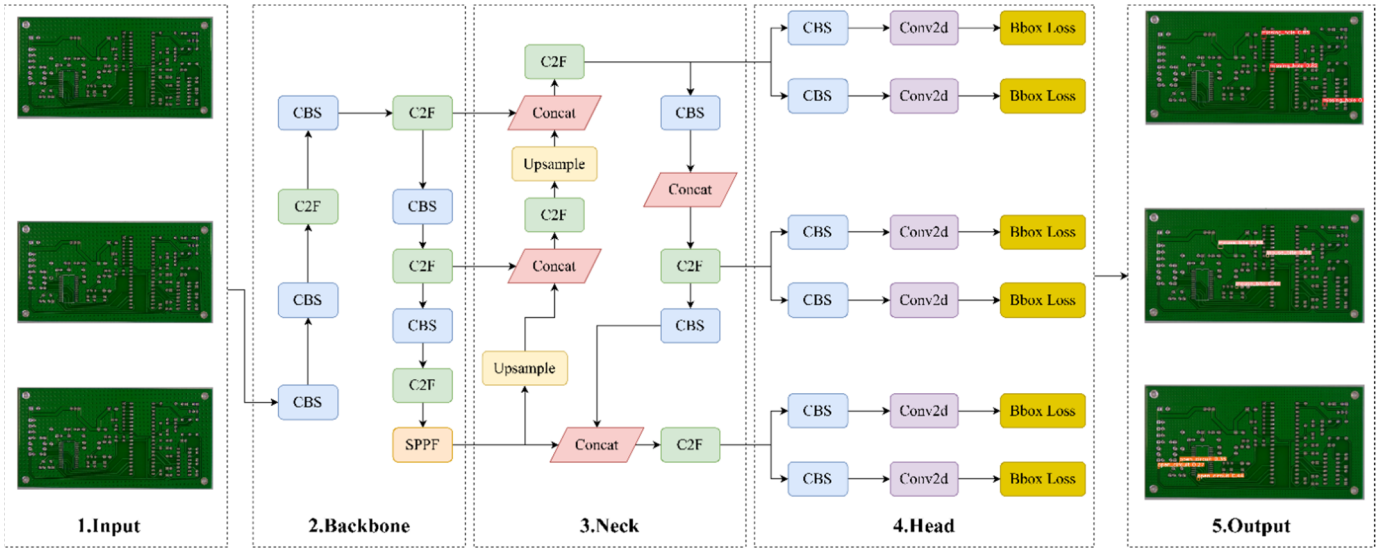


**Figure 2.** YOLOv8 network structure diagram

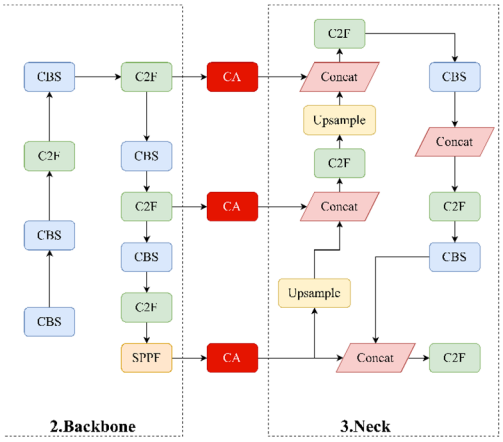


**Figure 3.** Picture of the dataset

The enhanced architecture of the network is seen in Figure 3. By inserting the CA module between the Backbone and Neck, the feature extraction capability between the two is fortified, enabling the feature extraction network to extract minute target defect features. This, in turn, facilitates the subsequent feature fusion process and enhances the accuracy of model detection.

# 3. Preparation of Experiments

## 3.1. Introduction to the Data Set

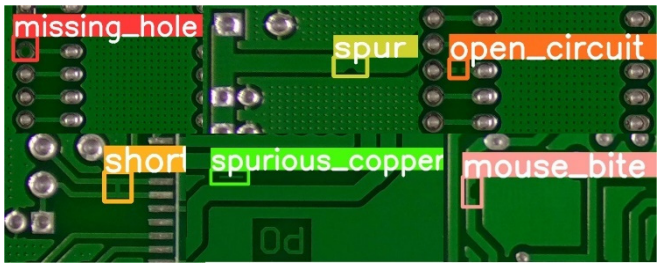


**Figure 4.** Picture of the dataset

As seen in Figure 4, the most prevalent surface defects on PCB boards include spur, mouse bite, short, spurious copper, missing hole, open circuit, and so on. A portion of the data utilized in this experiment is derived from a publicly available synthetic printed circuit board (PCB) dataset that was made available by Peking University. This dataset comprises 1386 pictures and six distinct types of faults. In a 7:2:1 ratio, the images are partitioned into a training set and a test set.

## 3.2. Indicators for Model Evaluation

The evaluation metrics contain parameters such as AP, mAP, Precision, Recall, F1-Measure (F1), Parameters, and FLOPs.

(1) Recall:

The ratio between the number of true case detections and the sum of true and false-negative case detections constitutes Recall. As an essential parameter for assessing the effectiveness of an object identification model, recall quantifies the model's capability to identify every object inside the image.

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

(2) Precision:

The ratio of true case detections to the sum of true and false positive case detections constitutes the accuracy rate. The precision rate is an extremely significant indicator that is utilized to assess the algorithm's Accuracy.

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

(3)F1:

The F1 score is calculated by summing the means of Precision and Recall. A solitary number encapsulates the object detection model's performance with respect to Recall and Precision. The F1 score is a metric that runs from 0 to 1, with a value of 1 signifying perfect model performance.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{10}$$

(4) AP:

AP is utilized to evaluate the target detection model's precision. The ratio of true case detections to the sum of true and false positive case detections constitutes accuracy.

$$AP = \int_0^1 p(r) \tag{11}$$

(5) Mean Average Precision(mAP):

Assuming the mAP represents the mean of the AP scores for every detected target category, it assesses the model's overall performance across all detected target categories. It is extensively employed as a standard against which various target detection models are evaluated.

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i \tag{12}$$

# 4. Analysis of Experimental Results

**Table 1.** Basic configuration of the local computer

| Computer Configuration | Specific parameters/versions |
|---|---|
| operating system | Ubuntu18.04 |
| CPU | AMD E5-2698B v3 |
| RAM | 15GB |
| GPU | P-48G |
| Python | 3.7 |
| PyTorch | 1.7.1 |
| Torchvision | 0.3.0 |
| CUDA | 11.1 |

The experimental setup described in this paper is an Ubuntu 18.04 system with the following components: 15GB of ROM, an AMD EPYC 7262 AMD E5-2698B v3 processor, a P4-8G GPU, and all models are written in PyTorch. The precise configuration of the local computer is detailed in Table 1.

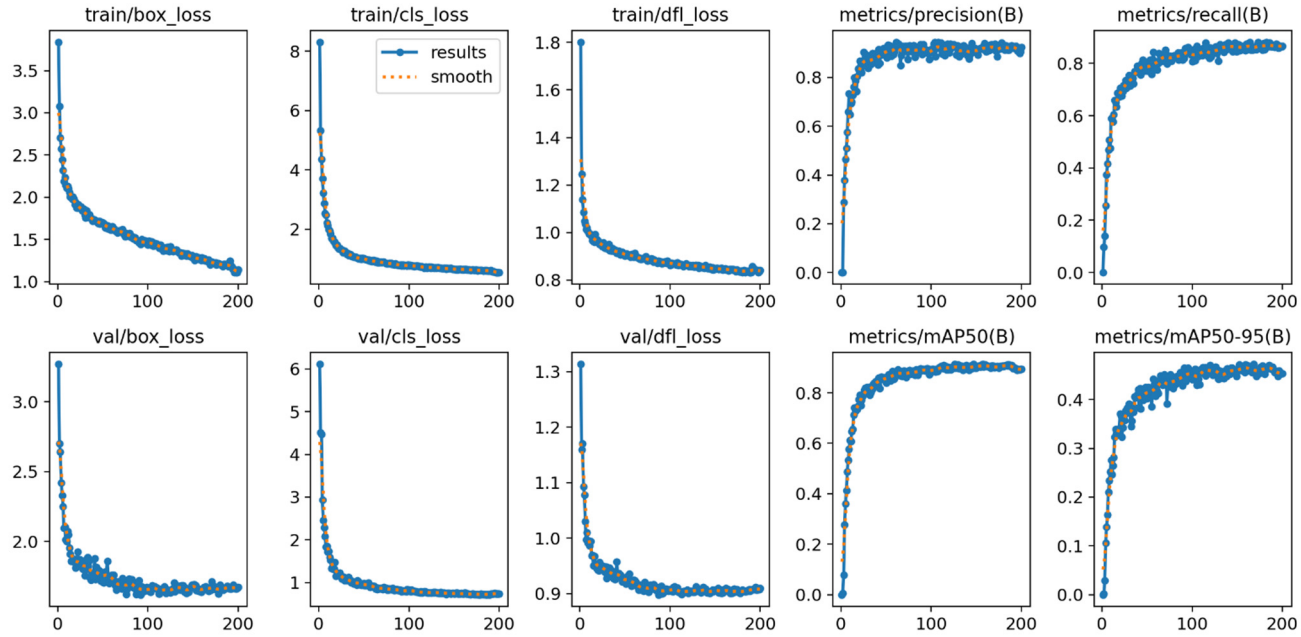


**Figure 5.** Model training results

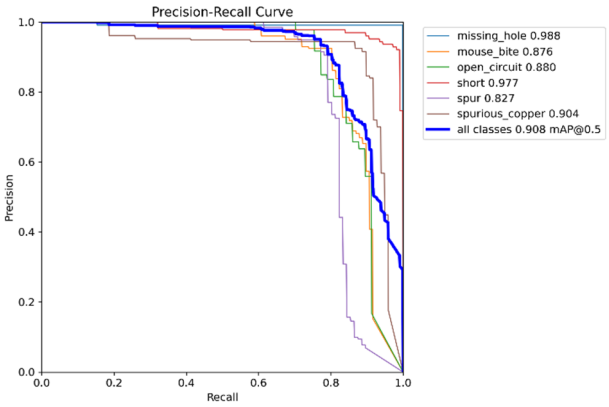


**Figure 6.** Diagram of model test results

The initial weight utilized by the model is the official weight YOLOv8.pt. The parameters for the input image configuration are as follows: 640 pixels by 640 pixels in dimensions, 8 pixels in batch size, 200 iterations as the maximum, and 0.001 as the starting learning rate. Between the fifty-first and one hundred fifty iterations. It then decays steadily, with a weight attenuation coefficient of 0.0005. As illustrated in Figure 5 and Figure 6, the training optimizer is SDG, the learning rate is maintained at 0.001, and all other parameters remain constant. This configuration facilitates the effective utilization of pre-existing data and reduces the duration of the training process. The outcomes of the model's training are as follows: mAP=90.8%, F1=0.97.

The performance of the YOLOv8s-CA model is compared to that of the YOLOv8s versions in this article. The experimental testing procedure involved the utilization of an identical test set, with all test parameters remaining constant. The model underwent testing, comparison, and analysis to validate the progress made on the enhanced model. The parameters of the AP, mAP, and F1 performance indices for various models are presented in Table 2.

**Table 2.** Model test result table

| Model type | mAP | Precision | Recall | F1 |
|---|---|---|---|---|
| YOLOv8-s | 84.2% | 100% | 95% | 0. 79 |
| YOLOv8-s-CA | 90.8% | 100% | 94% | 0.89 |

The enhanced YOLOv8-s-CA model has a 6.6 percent

increase in mAP value and a 0.1 percent gain in F1. This information is presented in Table 2. It demonstrates that the CA module optimizes the extraction of features from small targets, hence enhancing model precision and detection speed.

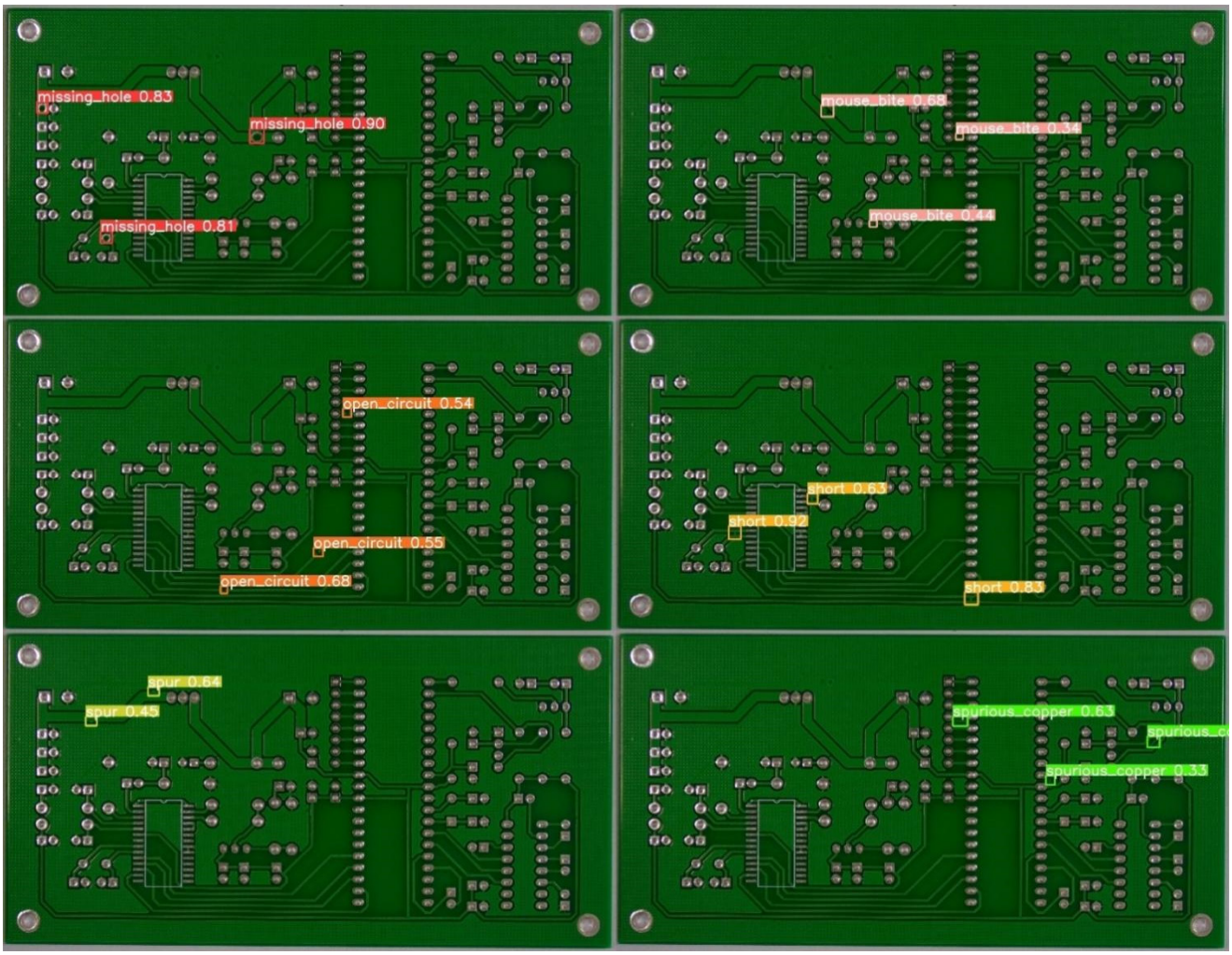


**Figure 7.** PCB detection results

Select some test set images to the YOLOv8s-CA algorithm model for real-time detection. The detection results are shown in Figure 7. The prediction box fits the area where the target is located. The improved model is more sensitive to small targets and has higher detection accuracy. The expected experimental requirements were met.

## 5. Conclusion

A strategy for enhancing the detection of surface defects on YOLOv8s PCB circuit boards is presented in this paper. A considerable variety of PCB board surface defects exist, each requiring a unique set of detection criteria. Complicating matters further, the lines of the PCB board resemble those of copper boards in color, which can readily obscure the visibility of burrs and residual copper. In order to resolve the issue, the CA attention mechanism is incorporated into the model to increase its sensitivity to spatial representation information and hence enhance its detection accuracy.

By means of several comparative experiments, it is demonstrated that the enhanced model can detect a variety of surface flaws on PCB circuit boards with precision and efficiency in real-time. Furthermore, it is straightforward to implement in small-scale industrial production. The detecting link is of the utmost importance in the industrial fabrication of printed circuit boards. As a result of the little data acquired at this moment, a comprehensive breakdown of different sorts of errors is not possible. In later endeavors, the algorithm will be enhanced regularly.

## References

[1] H. Suzuki, Official Gazette of the United States Patent and Trademark. Printed Circuit Board. U.S. Patent4,640,866, 16 March 1987.

[2] H. S. Cho, Samsung Electro Mechanics Co., Ltd. Official Gazette of the United States Patent and Trademark. Printed Circuit Board. U.S. Patent 8,159,824, 16 March 2012.

[3] R. Ding, L. Dai, G. Li, and H. Liu, “TDD-Net: A tiny defect detection network for printed circuit boardss,” CAAI Trans. Intell. Technol., vol. 4, no. 2, pp. 110-116, Jun. 2019.

[4] K. Goto, K. Kato, T. Saito, and H. Aizawa, “Adversarial autoencoder for detecting anomalies in soldered joints on printed circuit boardss,” Proc. SPIE, vol. 29, no. 4, 2020.

[5] I. C. Chen, R. C. Hwang, and H. C. Huang, “PCB Defect Detection Based on Deep Learning Algorithms,” Processes, 11 (3): 775,2023.

[6] J. Redmon, S. Divvala, R. Girshick, and Farhadi, “A. You Only Look Once: Unified, Real-Time Object Detections,” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.

[7] V. A. Adibhatla, “Applying deep learning to defect detection in printed circuit boards via a newest model of you-only-look-onces,” Mathematical Biosciences and Engineering,18(4): 4411-4428,2021.

[8] J. Tang, “PCB-YOLO: An Improved Detection Algorithm of PCB Surface Defects Based on YOLOv5s,” Sustainability, 15(7):5963,2023.

[9] B. Du, “YOLO-MBBi: PCB Surface Defect Detection Method Based on Enhanced YOLOv5s,” Electronics, 12(13): 2821, 2023.

[10] B. Hu and J. Wang, "Detection of PCB Surface Defects With Improved Faster-RCNN and Feature Pyramid Networks,” in IEEE Access, vol. 8, pp. 108335-108345, 2020.

[11] F. Wang, “Laser-induced thermography: An effective detection approach for multiple-type defects of printed circuit boards (PCBs) multilayer complex structures,” Measurement, Vol. 206, 2023.

[12] X. Y. Huo, “Development of a Real-time Printed Circuit Board (PCB) Visual Inspection System Using You Only Look Once (YOLO) and Fuzzy Logic Algorithms,” Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, Vol. 45, pp: 4139–4145.

[13] G. Wang, “UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios. Sensors,” 2023; 23(16):7190.

[14] F. Y. Zhou, “A small sample nonstandard gear surface defect detection method,” Measurement, Vol. 221,2023.

[15] J. Wang, “YOLOv7 Optimization Model Based on Attention Mechanism Applied in Dense Scenes,” Applied Sciences. 2023; 13(16):9173.