

# Deep Learning Based Defect Detection Research on Printed Circuit Boards

Qihang Yang

School of Computer Science and Engineering  
Xi'an Technological University  
Xi'an, 710021, China  
E-mail: gnay\_qh@163.com

Fan Yu

School of Computer Science and Engineering  
Xi'an Technological University  
Xi'an, 710021, China  
E-mail: yffshun@163.com

**Abstract**—As computer vision and deep learning detection techniques advance rapidly, their use in identifying defects has become more common across various industries. The significance of Printed Circuit Boards (PCBs) in contemporary electronic devices is undeniable, as they substantially influence the functionality and durability of these products. Thus, utilizing deep learning models for identifying flaws in Printed Circuit Boards (PCBs) is of particular importance. The focus of this study is primarily on examining PCB defect identification utilizing deep learning techniques. Firstly, it introduces the importance and development history of PCBs in the electronics and information industry. It then offers a comprehensive review of the existing research on conventional PCB defect detection approaches alongside methodologies grounded in deep learning. Following that, the structure of the YOLOv8 object detection model and its key technologies are elaborated. Lastly, the superior performance of YOLOv8 in PCB defect detection tasks is verified through comparative experiments. According to the evaluation metrics of the algorithm, the average detection accuracy reaches 92.3%, and the Frames Per Second (FPS) value reaches 157.2, meeting the accuracy requirements for PCB defect detection in the industrial domain.

**Keywords**—Deep Learning; Defect Detection; YOLOv8

## I. INTRODUCTION

The Printed Circuit Board (PCB), an essential element in electronic devices, is produced through processes that involve electronic chemistry, the industry is honored as the "mother of electronics". PCB plays an indispensable role in the electronic information industry. Widely used in a variety of fields, including but not limited to integrated circuits, artificial intelligence, medical equipment, aerospace and industrial equipment, PCB's main function is to connect the circuit components, so

as to facilitate the electronic equipment to achieve higher performance and efficiency. By connecting various electronic components and devices together in an orderly manner, PCB realizes the effective assembly of circuits, which is crucial for the proper functioning of electronic devices. In modern technology, the role of the PCB is not only to provide circuit connectivity, it also helps to optimize the stability, reliability and performance of electronic equipment. With the help of PCBs, device manufacturers are able to achieve higher performance standards and more sophisticated functionality, thus providing a better user experience. Therefore, PCB is considered one of the cornerstones in the field of electronics and is important for the development of modern society and technological advancement.

Modern electronic and electrical devices must rely on PCBs for electrical interconnections. Therefore, the quality of PCB boards is crucial for electronic devices and directly affects the success of the product. With the rapid development of emerging fields such as Internet of Things (IoT) technology, automotive electronics and 5G communications, the quality of PCB board design and production is vital for the efficiency and dependability of electronic products. Following the reform and opening up period, China's electronics sector has seen swift advancement, particularly in recent times, coinciding with the boom in the electronics industry, China's PCB manufacturing industry has continued to develop rapidly, with output value and production steadily ranking among the global leaders, making important contributions to the national economy and employment. With the continuous renewal of

electronic products, price competition continues to reshape the supply chain structure. With its industrial distribution, cost and market advantages, China has rapidly emerged as one of the most important PCB production bases in the world.

## II. RELATED WORKS

The PCB contains an extensive array of components, each with their intricate and varied characteristic details. During the production process, various uncertainties such as raw materials, production environment and manual operation often lead to various surface defects. Regular and prompt inspections are essential to maintaining production line functionality. As a result, there has been an ongoing investigation into PCB surface defect detection both domestically and internationally.

Currently, defect detection technologies are primarily divided into two approaches: conventional techniques and those based on deep learning. Next, the research status of these two directions will be described in detail.

Conventional methods for detecting defects in PCBs encompass manual examination, electrical measurement, and automated optical evaluation techniques. Manual inspection is one of the earliest approaches, which requires operators to use microscopes or magnifying glasses to discover various complex defects on the circuit board [1,2]. However, this approach is susceptible to subjective factors such as visual fatigue, which may lead to problems such as misdetection and missed detection. Simultaneously, the challenges associated with manual inspection increase because of the high cost of labor and reduced efficiency, compounded by the growing scale of integrated circuits and the intricacy of their wiring. Another method is electrical testing, which uses a probe instead of a needle bed, and a fast-moving electrical probe is placed in contact with the pins of the PCB for electrical measurement [3]. However, this approach requires the probes to be in contact with the board, which may damage the PCB surface and lead to unnecessary losses. Automated optical inspection, on the other hand, is limited by multi-sensor imaging, light source, field of view, and resolution. In real production

environments, debugging is complicated and lacks good portability. All of these traditional methods have limitations, the process is complex, and any one of the errors may lead to misdetection. A single inspection method can no longer meet the demand for efficient and rapid inspection of production lines.

The swift advancement of deep learning algorithms has sparked a surge in research focusing on the detection of PCB surface flaws using Convolutional Neural Networks (CNNs) [4]. Various researchers have introduced multiple deep learning-oriented techniques to address the issue of PCB defect detection. For instance, Ding and colleagues [5] have presented the Tiny Defect Detection Network (TDD-Net), which integrates the fundamental network of Faster R-CNN with Feature Pyramid Networks (FPN) [6] to enhance the precision in detecting PCB defects. In addition, Li et al [7] trained a mixture of Faster R-CNN and YOLOv2 models and integrated the detection results of the two models to achieve a high degree of precision. Hu and colleagues [8] have combined FPN with ResNet50 as the foundational network for Faster R-CNN, incorporating the ShuffleNetV2 framework to enhance the model's detection precision. Meanwhile, Tang and associates [9] employed a dual network alongside a pyramid pooling module (PPM) for defect identification, attaining high performance in detection. However, although the above algorithms have achieved better results in reducing false and missed detections, the intricate nature of the comprehensive model and its elevated time complexity pose challenges in fulfilling the real-time detection necessities within PCB manufacturing settings. This complexity may affect the application of the algorithms, especially in production sites where fast detection is required.

Yuan Li et al [10] integrated the Multi-Residual Attention Mechanism (MRHAM) into the YOLOv4 algorithmic model to enhance the ability of sensory field attention for defective target features. Concurrently, they employed the K-means++ clustering method to conduct an in-depth analysis of the PCB dataset, enhancing model robustness and ensuring rapid processing. Liao and colleagues [11] introduced a PCB defect

detection framework using YOLOv4-MN3, which streamlined and optimized the backbone network, feature fusion unit, and prediction module, diminishing the parameter count and resulting in a detection rate of 56.98 frames per second (FPS). In contrast, Wang and colleagues [12] presented a lightweight network for defect detection, YOLOX-MC-CA, which incorporated Coordinate Attention (CA) [13] and enhanced the CSPDarkNet backbone to accelerate detection speed, achieving satisfactory performance on the PCB open dataset. These research efforts meet the real-time criteria by streamlining the network architecture, thereby accelerating detection. Nevertheless, such simplification can potentially compromise the model's capacity to extract features from the input images. Therefore, further improving the detection efficiency while maintaining the detection accuracy is an important issue of concern for current scholars at home and abroad.

### III. ALGORITHMS MODEL

YOLOv8 was proposed by Ultralytics in January 2023 as an improved version of the

YOLOv5 algorithm model. Similar to YOLOv5, it does not have an associated paper at the moment, but its code has been open-sourced on the GitHub repository. YOLOv8 continues the overall architecture of the algorithm model since YOLOv4, as illustrated in Figure 1. The framework principally comprises three elements: the Backbone, which is the feature extraction network, the Neck, responsible for feature fusion, and the Head, the detection head component. Within this configuration, the Backbone commences the process by extracting attributes from the samples, resulting in the creation of feature maps at three distinct scales. The Neck integrates these three feature maps with surface and depth information, producing three new feature maps. The Head performs classification and regression on each sample point of the three new feature maps. The YOLOv8 algorithm model is categorized into N, S, M, L, and X versions based on the model's width and depth. In the following sections, a detailed analysis of YOLOv8 will be conducted, covering Backbone, Neck, Head, and the loss function.

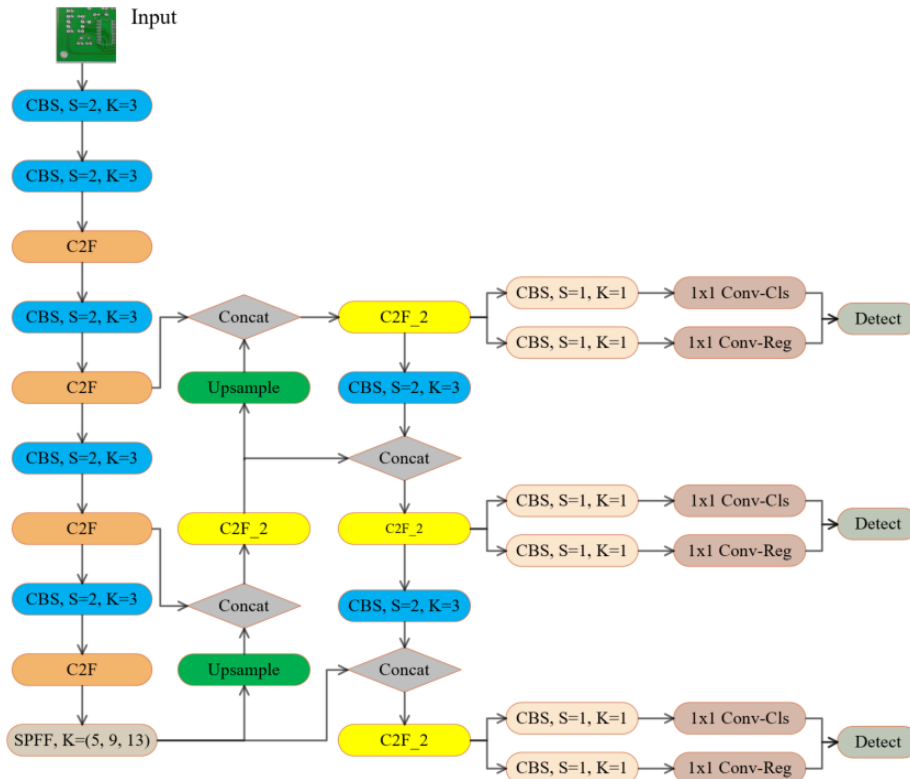


Figure 1. Diagram of YOLOv8 structure

*A. Feature Extraction Network*

The attribute extraction network of YOLOv8 represents an enhancement upon YOLOv5. Rather than employing the Focus component for parameter reduction, YOLOv8 utilizes a 2D convolution with a stride of 2 and a kernel size of 3 for expanding channels from the initial input. It also introduces a novel convolutional unit, C2f, which takes the place of the C3 unit in YOLOv5, while maintaining the SPPF pooling element.

The C2f module is still constructed based on the CSPNet (Cross Stage Partial Network) idea, and the architecture is depicted in Figure 2. Upon inputting the attributes into the C2f unit, the 1x1 convolution will be used for channel integration, and then the feature tensor will be sliced into two parts according to the channel Split, one of which will enter the Bottleneck block to further extract the features, and the other part will be spliced with the features processed by the Bottleneck block according to the channel. This configuration augments the ability of the convolutional neural network to extract features and minimizes the time spent on memory access. The Bottleneck is illustrated in Figure 3, which utilizes the residual idea, where the original input is convolved twice to extract features and then pointwise added to the original input. One part of the output of the Bottleneck continues to be used for the Bottleneck operation, and the other part is spliced with the half after Split for per-channel splicing. The advantage of the residual idea is that it not only preserves the basic features of the original input, but also avoids the problem of vanishing gradients.

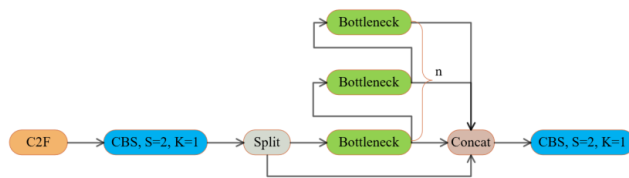


Figure 2. Schematic diagram of C2F module



Figure 3. Bottleneck Schematic Diagram

SPPF (Spatial Pyramid Pooling Fast) is an improved version of SPP, the structure is able to extract features at different scales of the object as SPP, enrich the feature information of the output layer of this feature, and present the same effect, while the time consumed is half of SPP, the structure is shown in Figure 4. The main process is: after the input features pass through the 1x1 convolutional integration channel, the output is copied in two copies, one for pooling kernel of 5 for maximum pooling, and the other is involved in the splicing with the output of the pooling layer. The pooled feature output is also copied into two copies, one to participate in the splicing of other pooled outputs, and the other to continue the maximum pooling, repeat this step three times, a total of four copies of the output features, four copies of the output features are spliced according to the channel, and the channel information is integrated using the 1x1 convolution.

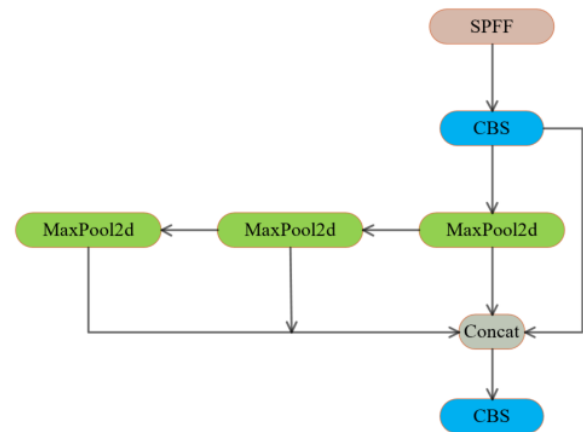


Figure 4. Schematic diagram of SPPF structure

*B. Feature Fusion Networks*

The Neck segment of YOLOv8 predominantly utilizes an enhanced version of the PANet (Path Aggregation Network) design, which builds upon the FPN (Feature Pyramid Networks). This structure involves upsampling the feature layers of smaller dimensions and fusing them with feature layers of larger dimensions, directly outputting the FPN to enrich the feature information contained in the large-dimensional feature maps. On the other hand, PANet further downsamples the fused large-dimensional features, merges them again with the

feature maps of smaller dimensions, enriching contextual information, and enhancing the expression capability of the smaller-dimensional features. In contrast to YOLOv5, YOLOv8 employs the C2f unit in place of the C3 unit, and integrates channel-wise before upsampling the feature maps of smaller dimensions by removing the 1x1 convolution.

### C. Detection Head

The Head component of YOLOv8 has transitioned from the initial coupled head design to the present prevalent decoupled head architecture. This decoupled head represents a standard design in object detection, tasked with deriving target location and class information from the detection network's feature map.

Specifically, decoupling the head involves separating the main part of the neural network model from the classifier part for training. The advantage of this design is the flexibility to modify and replace the classifier without altering the backbone network. Moreover, the decoupled head efficiently diminishes the quantity of parameters and the computational burden, which enhances the model's capacity for generalization and robustness, all while maintaining the backbone network architecture.

Within the YOLOv8 decoupled head structure, the prior Obj branch has been eliminated. Now, only separate branches for classification and regression persist. The regression branch employs an integral form derived from the Distribution Focal Loss concept. It is important to highlight that the channel counts in the classification and regression branches of the decoupled head may differ.

### D. Loss Function

The loss function is a critical element in the training of algorithmic models; an appropriately designed loss function can lead to quicker model convergence and enhanced robustness. The loss function utilized by YOLOv8 comprises a combination of classification loss VFL (Varifocal Loss) and regression loss CIOU + DFL (Distribution Focal Loss).

The classification loss VFL is shown in Equation 1, where  $q$  represents the overlap and concurrent alignment between the projected coordinate system and the actual coordinate system,  $p$  is the softmax output value of the category,  $\alpha$  considers the spectrum of values [0,1], and  $\gamma$  considers the spectrum of values [0,5]. Since there are too few positive samples during training, reducing the Loss contribution of negative samples makes the model more inclined to the training of high quality positive samples.

$$VFL(p, q) = \begin{cases} -q(q \log(p) + (1-q) \log(1-p)) & q > 0 \\ -\alpha p^\gamma \log(1-p) & q = 0 \end{cases} \quad (1)$$

The regression loss  $L_{reg}$  mainly calculated using the summation of the CIOU loss  $L_{CIOU}$  and the DFL loss is depicted in Equation 2, the  $L_{CIOU}$  mathematical expression is illustrated in Equation 3, and the DFL mathematical expression is illustrated in Equation 6.

$$L_{reg} = L_{CIOU} + DFL \quad (2)$$

$$L_{CIOU} = 1 - IOU(A, B) + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (3)$$

$$\alpha = \frac{v}{(1 - IOU) + v} \quad (4)$$

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

$$DFL(S_i, S_{i+1}) = -((y_{i+1} - y) \log(S_i) + (y - y_i) \log(S_{i+1})) \quad (6)$$

In Equation 3,  $IOU(A, B)$  signifies the degree of overlap and concurrent alignment between the actual and estimated coordinate systems,  $\rho^2(b, b^{gt})$  notes the geometric distance measured in Euclidean space between the central points of the forecasted and actual coordinate systems,  $c$  is

the diagonal distance of the outer rectangle containing the real and predicted target frames, and  $\alpha$  is the coefficient used for balancing the ratio as shown in Equation 4.  $v$  is the distance

between the height-width ratio  $\frac{w^{gt}}{h^{gt}}$  of the real

coordinate frame and the height-width ratio  $\frac{w}{h}$  of

the predicted coordinate frame, used to measure the height-width scale consistency as shown in Equation 5. In Equation 6,  $y$  denotes the true label value,  $y_i$  and  $y_{i+1}$  denote the two closest values of  $y$ , respectively, and  $S_i$  and  $S_{i+1}$  correspond to the probabilities of the two values.

#### IV. EXPERIMENTS

##### A. Experimental Environment

The experimental environment is shown in Table I

TABLE I. EXPERIMENTAL ENVIRONMENT

Experimental environment	Version
CPU	Intel Core i7-11800H
GPU	NVIDIA GeForce RTX307
Language	Python3.7
Deep Learning Framework	Pytorch1.11.0
CUDA	11.3.0
Compiler	Pycharm2021

##### B. Dataset

The dataset in question is a PCB (printed circuit board) defect collection made available by the Open Lab of Peking University. The types of defects are missing hole, mouse bite, open circuit, short circuit, spur, and spurious copper, and it contains a total of 11,361 images.

##### C. Evaluation Metrics

When assessing single-target detection models, it's customary to utilize metrics such as accuracy and recall to gauge the model's detection efficacy. The accuracy rate is defined as the ratio of correctly identified actual samples to the total number of samples, while the recall rate represents the ratio of correctly identified actual positive samples to the total number of actual positive

samples. The precise formulations for the accuracy rate  $P$  and the recall rate  $R$  are depicted in Equations 7 and 8, respectively.

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

Within this evaluative measure,  $TP$  signifies the count of positive samples rightly identified,  $FP$  denotes the quantity of negative samples erroneously labeled as positive, and  $FN$  indicates the number of positive samples mistakenly classified as negative. For a detection model, it is often desirable to have higher precision and recall, but it is often the case that a rise in one metric causes a fall in the other. In order to comprehensively assess the performance metric of a detection model, researchers introduce the P-R curve. The P-R curve is a curve that describes the change in the relationship between the model's accuracy and recall, and by determining the area beneath the curve, it can intuitively reflect the model's goodness or badness.

In the assessment of multi-target detection models, the metric often employed to gauge the comprehensive efficacy is the Mean Average Precision (mAP). This metric encapsulates the model's detection acuity across all categories. The mAP score is derived from the average of the Average Precision (AP) for each category. It mirrors the precision in detecting individual targets and is a function of the model's precision and recall rates. The underlying computational expressions are illustrated in Equations 9 and 10.

$$AP = \int_0^1 PRdR \quad (9)$$

$$mAP = \frac{1}{C} \sum_{c_i \in C} AP_{c_i} \quad (10)$$

D. Results

To validate the detection capabilities of the models presented in this study, a comparative analysis with contemporary mainstream object detection methodologies was conducted. The training of all models was carried out utilizing transfer learning techniques, with the COCO2017 dataset serving as the pre-training dataset. The outcomes of these experiments are documented in Table II.

TABLE II. EXPERIMENTAL RESULTS

Model	mAP(%)	FPS	Quantity of participants /M
Faster R-CNN	89.5	20.5	125.3
YOLOv5	85.3	79.4	42.7
YOLOv7	87.1	102.5	37.2
YOLOv8	92.3	157.2	28.5

The model introduced in this paper achieves superior overall performance regarding detection accuracy and speed, with a mAP of 92.3 and a detection rate of 157.2 frames per second (FPS), while only having 28.5 million parameters.

Although the two-stage approach, Faster R-CNN, offers higher detection accuracy due to its maximum input image resolution, its large parameter count and slower computation make it less practical for real-world applications. In contrast, single-stage models like YOLOv5 and YOLOv7 still trail behind YOLOv8 in both accuracy and speed. These findings are presented visually in Figures 5 and 6.

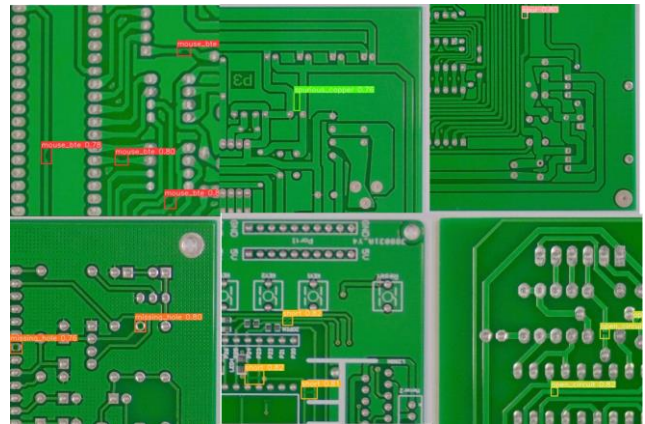


Figure 5. Pcb defect detection effect

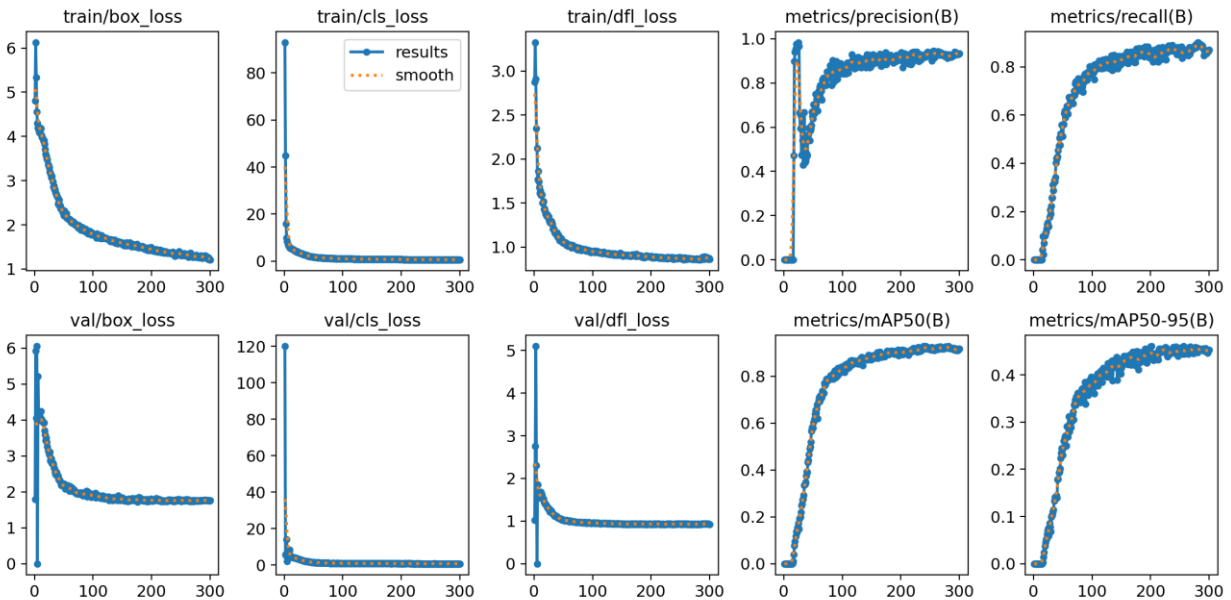


Figure 6. Loss, Precision, Recall, mPA0.5 and mAP0.5-0.95 curves

V. CONCLUSIONS

In this paper, we have explored the problem of PCB surface defect detection based on printed

circuit boards. Traditional PCB defect detection methods such as expert visualization or machine vision have more or less limitations, and with the evolution of deep learning target detection

algorithms in full swing, the application of deep learning methods to the PCB defect detection problem is also a general trend. Due to the industrial production of PCB surface defects there are difficult to detect the problem, this paper carries out a series of analyses of the existing detection model, and finally in the experimental environment under the same conditions, for the detection model of YOLOv8 and the other mainstream detection models in the industry in the detection of the average accuracy of the average value, detection speed, model complexity and other indicators of the assessment of a comprehensive comparison of the experiments. Furthermore, the experimental results suggests that the YOLOv8 detection algorithm exhibits a significant enhancement in the mentioned metrics, demonstrating its relevance for identifying defects in electronic products. This underscores its valuable research potential within the realm of object detection.

#### REFERENCES

- [1] Li Y, Li S. Defect detection of bare printed circuit boards based on gradient direction information entropy and uniform local binary patterns [J]. *Circuit World*, 2017, 43(4): 145-151.
- [2] Chomsuwan K, Yamada S, Iwahara M. Improvement on defect detection performance of PCB inspection based on ECT technique with multi-SV-GMR sensor [J]. *IEEE Transactions on magnetics*, 2007, 43(6): 2394-2396.
- [3] Verdingovas V, Jellesen M S, Ambat R. Colorimetric visualization of tin corrosion: A method for early stage corrosion detection on printed circuit boards [J]. *Microelectronics Reliability*, 2017, 73: 158-166.
- [4] Yang J, Li S, Wang Z, et al. Using deep learning to detect defects in manufacturing: a comprehensive survey and current challenges [J]. *Materials*, 2020, 13(24): 5755.
- [5] Ding R, Dai L, Li G, et al. TDD - net: a tiny defect detection network for printed circuit boards [J]. *CAAI Transactions on Intelligence Technology*, 2019, 4(2): 110-116.
- [6] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.
- [7] Li Y T, Kuo P, Guo J I. Automatic industry PCB board DIP process defect detection with deep ensemble method [C]//2020 IEEE 29th International Symposium on Industrial Electronics (ISIE). IEEE, 2020: 453-459.
- [8] Hu B, Wang J. Detection of PCB surface defects with improved faster-RCNN and feature pyramid network [J]. *Ieee Access*, 2020, 8: 108335-108345.
- [9] Tang S, He F, Huang X, et al. Online PCB defect detector on a new PCB defect dataset[J]. *arXiv preprint arXiv:1902.06197*, 2019.
- [10] Li Yuan. Research and application of target detection method based on improved YOLOv4 [D]. Chongqing: Chongqing University of Posts and Telecommunications, 2021.
- [11] Liao X, Lv S, Li D, et al. Yolov4-mn3 for pcb surface defect detection [J]. *Applied Sciences*, 2021, 11(24): 11701.
- [12] Xuan W, Jian-She G, Bo-Jie H, et al. A lightweight modified YOLOX network using coordinate attention mechanism for PCB surface defect detection [J]. *IEEE Sensors Journal*, 2022, 22(21): 20910-20920.
- [13] Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 13713-13722.