2023 5th International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI) | 979-8-3503-5993-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/MLBDBI60823.2023.10482094

# PCB surface defect detection based on improved YOLOv7-tiny

Hang Wu
School of Mechanical Engineering
Jiangsu University of Technology
Changzhou, 213001, China
wh19972023@163.com

* Mingyue Tang
School of Mechanical Engineering
Jiangsu University of Technology
Changzhou, 213001, China
aqwerdf88@163.com

***Abstract*—To tackle the challenges of low detection accuracy and complex detection difficulty associated with surface defects on PCB, this paper introduces a surface defect detection algorithm for PCB. First, CARAFE is used as an upsampling method for the YOLOv7-tiny algorithm to expand the sensing field and improve the perception of PCB surface defects. Secondly, Global Attention Mechanism is introduced at the head output network front end of the network model to enhance the extraction capability of PCB surface features, making the algorithm focus more on important feature regions. Finally, the CIoU loss function is modified to Focal-EIoU, which solves the imbalance problem of training samples and enhances the target localization accuracy and regression accuracy. Through the experiments, that the mAP of the proposed new algorithm has risen by 5.1%, and the FPS reaches 99. The improved algorithm significantly outperforms other mainstream algorithms in terms of comprehensive performance.**

***Keywords-component; YOLOv7-tiny; PCB surface defect detection; CARAFE; Global Attention Mechanism; Focal-EIoU***

## I. INTRODUCTION

Printed Circuit Board (PCB) finds extensive application in artificial intelligence product development, autonomous driving, aerospace and other fields. However, Given the intricate nature of PCB manufacturing and the stringent demands for precision, PCB is susceptible to defects due to factors such as the production environment, equipment, operator misuse, and technical failures. Therefore, it is crucial to detect defects in PCB, such as can reduce the safety hazards caused by quality problems as well as reduce economic losses.

The PCB was initially inspected using manual visual observation and electrical inspection methods. However, manual visual inspection is inefficient, time-consuming, and prone to omissions and errors. Electrical inspection methods, in turn, have limitations that prevent them from detecting all types of defects, and there is a risk of scratches when touching the PCB with specialized instruments.

In recent years, researchers have used machine vision-centered algorithms for PCB surface defect detection. Some of these methods may adversely affect inspection accuracy during alignment with standard templates due to various factors such as size and brightness. In addition, some of the inspection processes are intricate and time-intensive, making it difficult to improve the inspection efficiency in the actual complex PCB production industrial environment.

In the realm of artificial intelligence target detection, researchers are increasingly adopting deep learning algorithms on a wide scale. One-stage target detection algorithms represented by the SSD [1] and YOLO [2-4] series are one class, while two-stage target detection algorithms represented by R-CNN [5], Fast R-CNN [6] and Faster R-CNN [7] are one class. Wu et al [8] referenced MobileNetV3 as the feature extraction network and Inceptionv3 as the detection network in the YOLOv4 algorithm, which effectively improves the detection of PCB. However, the algorithm model volume of 53.2MB is too large, which is not suitable for mobile deployment applications in practical engineering.Li et al [9] based on Faster R-CNN and FPN as the architecture, the feature extraction backbone network introduces the SE module, introduces an enhanced bottom-up structure, and finally uses ROI Align instead of RoI Pooling, the experimental results indicate an enhancement in the detection efficiency, but it will make the algorithm more complex, the computational cost is getting higher and higher, and it will drastically reduce the detection speed.Li et al [10] used DBSCAN+k-means clustering algorithm on the basis of YOLOv3 algorithm to generate the aiming frame, increased the residual unit to improve the network's ability to extract shallow features, and introduced the SE Block module to improve the feature fusion structure to improve the detection ability of PCB dataset, but its mAP is only 87.6%, which is low.Although the above algorithm has made some optimization improvements in PCB defect detection, there are still deficiencies in the detection ability, such as large model volume, low detection accuracy, high network complexity, slow detection speed and other problems.

To tackle the previously mentioned issues, this paper presents a PCB defect detection approach leveraging the enhanced YOLOv7-tiny.YOLOv7 is introduced by Wang et al [11] in July 2022, and there are several model versions, such as YOLOv7-X, YOLOv7-w6, etc. YOLOv7-tiny has the advantages of fewer parameters, simple network structure , low requirements on experimental equipment and other advantages. Therefore, the YOLOv7-tiny algorithm is chosen to be improved in this paper.This paper makes three improvements, CARAFE is used as an upsampling method for the YOLOv7-tiny algorithm to expand the sensory field of the algorithm. This improvement helps to increase the model's ability to sense PCB surface defects. The GAM is introduced at the initial stage of the model head's output network to enhance the algorithm's focus on the defect feature regions. which can enhance the algorithm's feature extraction ability for PCB surface defects. Substituting the original CIoU loss function with the Focal-EIoU loss function effectively addresses the issue of imbalanced samples during training, improves the accuracy of target localization, and effectively improves the regression accuracy.The improved algorithm maintains the need for real-time detection with almost no increase in computational complexity, and the detection accuracy is greatly improved. The improved YOLOv7-tiny network structure is shown in Figure 1.

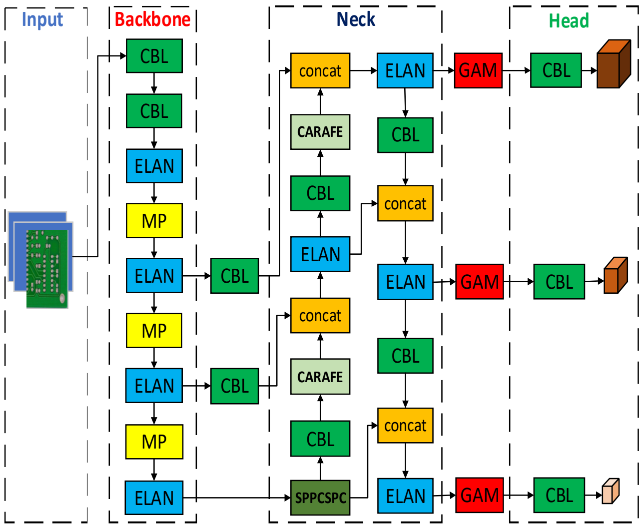



Figure 1. Diagram of the improved YOLOv7-tiny network structure

## II. Improvement of YOLOv7-tiny

### A. CARAFE

The nearest neighbor interpolation algorithm used as upsampling in YOLOv7-tiny. The algorithm only considers the spatial location of the pixel point, resulting in a small perceptual range. As a result, it fails to capture sufficient information from the feature map of small targets.

In this paper, the CARAFE operator [12] is used instead of the nearest neighbor interpolation algorithm. It can aggregate contextual information over a larger range of perceptual domains, adaptively perceive content, and obtain more semantic information, and more accurately capture the defective features on the PCB surface. As shown in Figure 2, It has a kernel prediction module and a content-aware reorganization module.

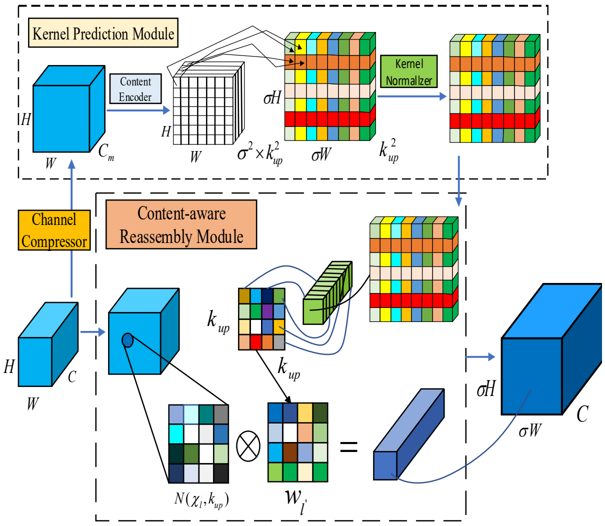



Figure 2. CARAFE module structure

### B. Global Attention Mechanism

In this paper, the PCB dataset has a wide variety of defects, different shapes, and only occupies a small part of the image, which is easy to interfere with the detection. To enhance the algorithm's capability to extract features for defective areas on PCB, three GAM modules are introduced at the network front end of the algorithm.

The Global Attention Mechanism (GAM) [13] has the potential to boost the performance of deep neural networks by minimizing information loss and amplifying the interaction between global dimensional features. This leads to an improvement in model expressiveness and learning. Figure 3 depicts the module structure of GAM.

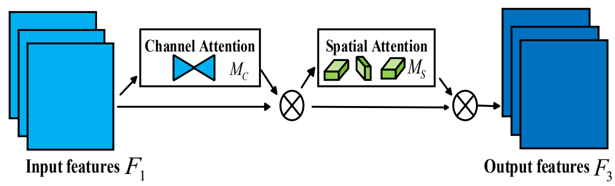



Figure 3. GAM module structure

Where $M_C$ is the channel attention module and $M_S$ is the spatial attention module. The given input feature is denoted as $F_1$ ,the intermediate state is denoted as $F_2$, and the final output feature is $F_3$. The relationship between $F_1$ and $F_2$ and $F_3$ is shown in equation (1) and equation (2):

$$F_2 = M_C(F_1) \otimes F_1 \tag{1}$$

$$F_3 = M_S(F_2) \otimes F_2 \tag{2}$$

### C. Focal-EIoU loss function

The CIoU loss function in the YOLOv7-tiny algorithm tends to lead to uneven training samples, unstable convergence speed, and degradation of regression accuracy. Therefore, using the Focal-EIoU [14] loss function instead of the CIoU function, solve the problem of unbalanced training samples, improve the accuracy of target localization, and improve the regression accuracy. Its formula is defined as follows:

$$L_{Focal-EIoU} = IoU^{\gamma} L_{EIoU} \tag{3}$$

$$L_{EIoU} = 1 - IoU + \frac{\rho^2(b,b^{gt})}{c^2} + \frac{\rho^2(w,w^{gt})}{c_w^2} + \frac{\rho^2(h,h^{gt})}{c_h^2} \tag{4}$$

where IoU denotes the intersection and merger ratio, $\gamma$ denotes a parameter used to regulate the extent of outlier suppression, $L_{EIoU}$ is denoted as the EIoU loss function.

## III. EXPERIMENT

### A. Experimental environment and parameter

The experimental platform in this paper utilizes Ubuntu 18.04 as the operating system. The hardware environment consists of an NVIDIA GeForce RTX 3090 graphics card with 24 GB of video memory and an Intel(R) Xeon(R) Platinum 8255C CPU processor, and the main programming language is Python 3.8, and PyTorch is used as the deep learning framework.

The size of the input image is uniformly set to 640×640, the model training batch size is set to 16, and the Epoch is set to 300.

### B. Experimental data set

The dataset employed for the experiments in this paper is a publicly available dataset provided by the Open Laboratory for Intelligent Robotics [15] at Peking University. There are 693 original images in this PCB dataset, and there are a total of six common defects, such as: missing_hole, mouse_bite, open_circuit, short, spur, and spurious_copper.

Because of the small dataset used, to prevent model overfitting, the dataset size is effectively increased through random rotations, cropping, and scaling of the original images. This augmentation enhances the model's generalization ability and robustness. Ultimately, the dataset expands to 15857 images, which are then partitioned into training and testing sets at a ratio of 7:3.

### C. Evaluation metrics

By using the mean of average precision across all categories (mAP), floating point calculations (FLOPs) for model complexity, Model Size, and the number of frames per unit of time (FPS) detected by the images in the model as evaluation metrics. The mAP formula is defined as follows:

$$P = \frac{TP}{FP + TP} \tag{5}$$

$$R = \frac{TP}{FN + TP} \tag{6}$$

$$AP = \int_0^1 P(R)dR \tag{7}$$

$$mAP = \frac{\sum_{i=1}^{C} AP_i}{C} \tag{8}$$

where $P$ denotes precision; $R$ denotes recall; $TP$ denotes the number of correctly predicted positive samples; $FP$ denotes the number of incorrectly predicted positive samples; $FN$ denotes the number of incorrectly predicted negative samples; $AP$ denotes the average precision of each target category; and $C$ is the total number of target categories.

### D. Ablation experiment

As shown in Table I, this paper, several rounds of ablation experiments are conducted to showcase the algorithm's efficacy.

TABLE I. ABLATION EXPERIMENT

| | CARAFE | GAM | Focal-EIoU | mAP% | Model Size/MB | FLOPs(G) |
|---|---|---|---|---|---|---|
| A | × | × | × | 90.4 | 11.7 | 13.1 |
| B | √ | × | × | 92.5 | 11.8 | 13.2 |
| C | √ | √ | × | 93.6 | 12.3 | 14.7 |
| D | √ | √ | √ | 95.5 | 12.3 | 14.7 |

As indicated in Table 1, the original YOLOv7-tiny algorithm has a mAP of 90.4%. When only the nearest neighbor interpolation upsampling was changed to CARAFE, the model's receptive field is increased and the mAP improved by 2.1% over the original algorithm. Subsequently, three GAMs are introduced at the front end of the model's header output network, which caused the model to focus more on important feature regions, and the mAP improved by 3.2% over the original algorithm. The final improved algorithm is after adding CARAFE and GAM and modifying the CIoU loss function to Focal-EIoU loss function to solve the training sample imbalance problem, which improves the accuracy of target localization and enhances the regression accuracy.The improved algorithm detects a mAP of 95.5%, which is a significant improvement of 5.1% over the original algorithm.

### E. comparison experiment

Table 2 shows that the improved algorithm is compared with other deep learning algorithms, using Model Size, mAP, and FPS as evaluation metrics.

TABLE II. ALGORITHM COMPARISON EXPERIMENT

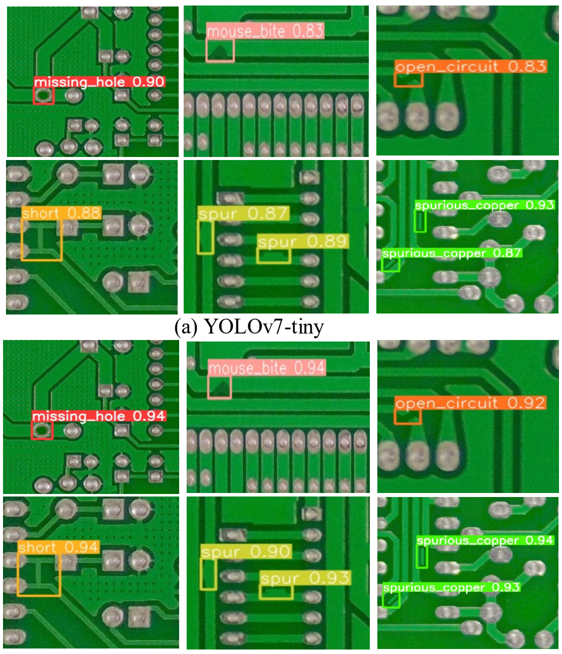
| Algorithm | Model Size/MB | mAP% | FPS |
|---|---|---|---|
| SSD | 93.1 | 85.5 | 25 |
| Faster R-CNN | 103.0 | 88.6 | 27 |
| YOLOv6s | 38.6 | 92.0 | 90 |
| YOLOv7 | 71.4 | 91.2 | 75 |
| YOLOv7-tiny | 11.7 | 90.4 | 107 |
| ours | 12.3 | 95.5 | 99 |

Table II shows that the improved YOLOv7-tiny algorithm in has good overall performance.The two algorithms, Faster R-CNN and SSD, have larger Model Size, poorer FPS, and the lowest recognition accuracies, which are unsuitable for mobile deployment. Although YOLOv6s algorithm and YOLOv7 algorithm have high detection accuracy, the performance of Model Size and FPS is not as good as the improved YOLOv7-tiny algorithm.

Therefore, the improved algorithm has the highest mAP on the PCB dataset and outperforms other detection algorithms. Although the FPS decreases slightly, the FPS reaches 99, which still meets the requirement of good real-time PCB detection. In addition, the algorithm is small in size and low in complexity, and can be quickly deployed in industrial inspection equipment.

### *F. Visual analysis*

As to visualize the detection effect before and after the improvement of the algorithm, six types of defect pictures are selected and their detection results before and after the improvement of the algorithm are enlarged and displayed in Figure 4.




(a) YOLOv7-tiny

(b) Improved YOLOv7-tiny

Figure 4. Comparison of Algorithm Improvement Detection Results

Through the comparative detection demonstration before and after the algorithm improvement in Figure 4, it can be intuitively observed that the improved algorithm has obviousness in the detection effect, and can identify the defective targets of PCB with high accuracy under different complex background environments.

## IV. Conclusions

This paper presents a novel approach for detecting PCB defects, utilizing an enhanced variant of the YOLOv7-tiny algorithm. The proposed method solves the problems of low accuracy and complexity of PCB surface inspection. The nearest-neighbor interpolation upsampling module, which is replaced by CARAFE, is used to expand the sensing field of the model, which improves the ability to perceive defects on the PCB surface. Incorporating the GAM at the front end of the model's output network significantly enhances its feature extraction capabilities from the PCB surface. This allows the model to focus more on important feature regions, leading to improved detection accuracy. At the same time, the CIoU loss function in the model is replaced with the Focal-EIoU loss function, which solves the imbalance problem of training samples, enhances the accuracy of target localization, and enhances the precision of regression.These three modular improvements are able to combine the advantages of each, and the optimized YOLOv7-tiny algorithm demonstrates excellent performance in terms of Model Size, FPS and mAP.

## References


[1] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21-37.

[2] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.

[3] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7263-7271.

[4] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv: 1804.02767, 2018.

[5] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.

[6] Girshick R. Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1440-1448.

[7] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28.

[8] Wu J G, Cheng Y, Shao J, et al. A defect detection method for PCB based on the improved YOLOv4[J]. Chinese Journal of Scientific Instrument, 2021, 42(10): 171-178.

[9] Li D, Fu S L, Zhang Q J, et al. An improved PCB defect detector based on feature pyramid networks[C]//Proceedings of the 2020 4th International Conference on Computer Science and Artificial Intelligence. 2020: 233-239.

[10] Li W, Li X, Yan H. PCB defect detection based on improved YOLOv3[J]. Electron. Opt. Control, 2022, 29(4): 106-111.

[11] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7464-7475.

[12] Wang J, Chen K, Xu R, et al. Carafe: Content-aware reassembly of features[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2019: 3007-3016.

[13] Liu Y, Shao Z, Hoffmann N. Global attention mechanism: Retain information to enhance channel-spatial interactions[J]. arXiv preprint arXiv:2112.05561, 2021.

[14] Zhang Y F, Ren W, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.

[15] 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.