Journal of Advanced Research in Applied Sciences and Engineering Technology 31, Issue 3 (2023) 249-264

SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

CSSD-YOLO: A Modified YOLOv5 Method for Solder Joint Defect Detection

Li Ang[1,2], Raseeda Hamzah[3,*], Siti Khatijah Nor Abdul Rahim[1], Gao Yousheng[1,2], Khyrina Airin Fariza Abu Samah[3]

[1] College of Computing, Informatics and Media, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor Malaysia
[2] College of Information Engineering, Jiujiang Vocational University, Jiu Jiang, Jiang Xi, China
[3] Computing Sciences Studies, College of Computing, Informatics and Media, Universiti Teknologi MARA (UiTM) Melaka Branch, Malaysia

ARTICLE INFO

Article history:
Received 25 April 2023
Received in revised form 27 July 2023
Accepted 2 August 2023
Available online 16 August 2023

Keywords:
Deep Learning; solder joint defect; object detection; feature extraction; attention mechanism

ABSTRACT

Surface Mount Technology (SMT) pin solder joint defects are hard to detect since the joints are smaller and denser and have high similarity between defect samples and standard samples in solder joint images. We propose an improved YOLOv5 defect detection algorithm embedding Cascade Shuffle Space to Depth (CSSD), Coordinate Attention (CA) mechanism module, and K-means++ algorithm. The proposed improved Yolov5 significantly impacts the loss and model parameter reduction and higher positioning precision of the defect location on the disk. The optimum anchor box produces better clustering and stability. Compared with the original YOLOv5 under the same test conditions, the method in this paper improves the precision by 12.2%, recall by 9.4%, mAP by 9.0%, and model parameters reduced by 1.3M. In conclusion, the experimental results show that the algorithm proposed in this paper has a better detection effect and a smaller parameter scale. It also can better meet the defect detection and model deployment in the actual industrial production environment.

1. Introduction

With the development of computer technology, artificial intelligence, 5G communication, and other emerging technologies, electronic products have become an indispensable part of social life [1,2]. The rapid development of emerging technologies puts forward higher requirements for the performance and quality of electronic products. Surface Mount Technology (SMT) chips are widely used in various electronic devices and are essential in electronic circuits. In the production process, SMT chip pins are soldered onto the PCB by soldering tin, which is prone to produce some defects during soldering, such as insufficient defects and pin shifting defects [3], as shown in Figure 1. These defects will cause electronic equipment to fail to operate normally, so detecting solder joint defects of SMT chips is crucial. Traditional solder joint detection adopts manual detection, which is easy to reduce the detection rate due to the visual fatigue of workers, and the labor cost is high. There are two significant difficulties in SMT chip pin solder joint defect detection.

* *Corresponding author.*
E-mail address: raseeda@uitm.edu.my

https://doi.org/10.37934/araset.31.3.249264

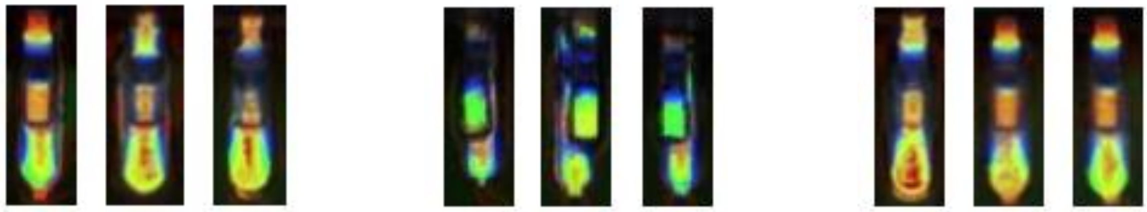
i. Compared with component solder joints, SMT chip pin solder joints are smaller and denser. Smaller defect areas and lower image resolution make detection more difficult.
ii. No significant difference exists between qualified and defective solder joints. Based on these difficulties, this paper proposes a solder joint defect detection method based on a lightweight deep learning object detection algorithm, which has a good detection effect on SMT chip pin solder joint defects.

The machine vision system embedding machine vision technology is widely used in various industrial product defect detection scenarios because of its high detection efficiency and the advantages of reducing labor costs [4-6]. The detection of solder joint defects based on machine vision can be divided into three categories: the feature-based method [7], the statistic-based method [8], and the deep-learning-based method [9]. The deep learning method has been used extensively in this research area, especially in developing a lightweight deep neural network (DNN) model.

The types of lightweight DNN models can be divided into two stages and single-stage. Although the two-stage object detection represented by Faster R-CNN performs well in detection accuracy, its model is relatively complex and has poor real-time performance. It is hard for a two stages architecture to meet the speed requirements in industrial scenarios. Zhang *et al.,* [10] applied a Faster R-CNN algorithm to inspect solder joints in the connectors. Based on the Faster R-CNN algorithm, ResNet-101 is used to replace VGG-16. The method has produced better accuracy; however, the processing speed is low. Ding *et al.,* [11] proposed a tiny defect detection network (TDD-Net) based on Faster R-CNN to improve the performance of PCB defect detection. The original Faster R-CNN network structure was improved. The online hard sample mining was used to effectively utilize the data information to realize the high-precision detection of PCB defects. Han *et al.,* [12] used Mobile-Net as the primary feature extraction network of Tiny-YOLOv3 for aerospace electronic solder joint defects, which improved the detection accuracy while maintaining the detection speed. However, the overall detection accuracy still needs to be improved. The SMT pin solder joint defect detection method based on deep learning is rare.

In 2018, [13] proved that a cascade-based deep learning algorithm reduces the training's memory usage and time requirements compared to traditional end–end backpropagation. It circumvents the vanishing gradient problem by learning feature representations that have an increased correlation with the output on every layer. Cai *et al.,* [14] designed a cascade deep learning network framework with three convolution neural networks. Through the trained CNN framework, the potential features of IC solder joints are adaptively extracted to achieve an accurate prediction. [15] implements cascade neural networks to associate between typical defect characteristics and the substation types of equipment to avoid manual selection for preliminary image annotation. Ye *et al.,* [16] propose a cascade neural network for diagnosing malignant tumors in histopathological sections of common eyelid tumors, improving detection accuracy. Tang *et al.,* [17] proposed a spatiotemporal cascade neural network for video salient object detection, a cascade of two complete convolution networks to evaluate visual saliency from spatial and temporal cues, resulting in the optimal video saliency prediction. Dong *et al.,* [18] proposed a cascade neural network for object detection in high-resolution remote sensing images, which combines the first-order statistical features of samples with the BP neural network model to improve the recognition effect of detection methods. Although cascade-based deep learning algorithms have shown efficiency, there is still room for improvement.

YOLO is a single-stage object detection algorithm combining the two stages of generating candidate regions and detection. It only needs one step to get the position of the object to be detected and can perform faster detection speed. YOLO model uses adaptive anchor frame calculation and multi-semantic fusion detection mechanism to integrate rich high-level semantic information quickly and effectively with low-level location information to achieve rapid object detection [19]. YOLOv5 has the advantages of a small model, fast speed, and high accuracy [20]. In order to meet the need for real-time detection of industrial production lines, this paper selects the YOLOv5 algorithm to improve detection accuracy.



(a) Insufficient defect (b) Foot shifting defect (c) Qualified samples

Fig. 1. Insufficient defect(a), foot shifting defect(b) and qualified(c) solder joint samples in dataset.

YOLOv5 has a relatively serious phenomenon of missing detection in the detection of small target defects. The reason is that in the process of feature extraction, the convolution kernel performs the convolution operation in two strides resulting in omittion of some important features such as small target and its important medium, shallow texture and contour information. Reducing the convolution stride can retain more effective features without being missed, but it will greatly increase the number of the model parameters. F. Chollet [21] proposes a depthwise separable convolution, which consists of depthwise convolution and pointwise convolution. The input feature map is grouped by depthwise convolution, the number of input channels and the output channels are kept consistent. Pointwise convolution through 1×1 normal convolution changes the number of input channels. It can reduce the number of parameters and the computational complexity of the model, and improve the efficiency of the use of convolution kernel parameters. Zhang *et al.,* [22] proposed a shuffle operation. The shuffle operation is a feature fusion method. This method shuffles and rearranges the features from the convolution of different channels to help information flow between different channels. It is proved that the channel shuffle operation is more conducive to the performance improvement of the model, and further verifies the importance of feature fusion between different channels. Sajjadi *et al.,* [23] proposed a space to depth transformation idea for video super-resolution tasks, extracts shifted low resolution grids from the image and places them into the channel dimension. This method can be used for down-sampling. Compared with ordinary down-sampling, it can retain more features and realize down-sampling operation with less computational cost.

Based on the above ideas, this paper proposes a Cascade Shuffle Space to Depth (CSSD) module to enhance the feature extraction ability of YOLOv5, reduce the loss of effective features in the feature extraction process. It can extract features more comprehensively and control the surge of model parameters.

The major innovations and contributions of this research are as follows:

i. In order to improve the detection ability of YOLOv5 for solder joint defects, a CSSD module is proposed, which integrates into the backbone network and neck network of YOLOv5. While improving the feature extraction ability of the model for defective targets, the model parameters are reduced.
ii. In order to enhance the precise positioning ability of the model to the defect location, the CA attention mechanism module is introduced to enhance the receptive field, and three CA module embedding structures are proposed to improve the defect detection accuracy.

iii. In order to obtain priori anchor boxes that is more consistent with the solder joint defect and improve the efficiency of network training, K-means++ algorithm is used to analyze the solder joint defect dataset, overcomes the poor clustering effect and stability caused by the random selection of initial values in K-means algorithm, and reduces the error caused by the size of the anchor boxes.

The rest of the paper is arranged as follows: Section 2 introduces the basic network structure of the original algorithm YOLOv5; introduces the improvement strategy of CSSD-YOLO in detail, including the composition of CSSD module and the embedding of CA attention mechanism module; Section 3 introduces data collection, experimental environment, comparative experiment, ablation experiment results and analysis; Section 4 summarizes the work and puts forward the follow-up work and improvement direction.

2. Methodology

2.1 YOLOv5

This paper comprehensively considers the model size, parameter quantity, etc. YOLOv5s model is selected as the solder joint defect detection network. YOLOv5 network structure is composed of Backbone network, Neck feature fusion network and detection head [24]. Its network structure is shown in Figure 2.

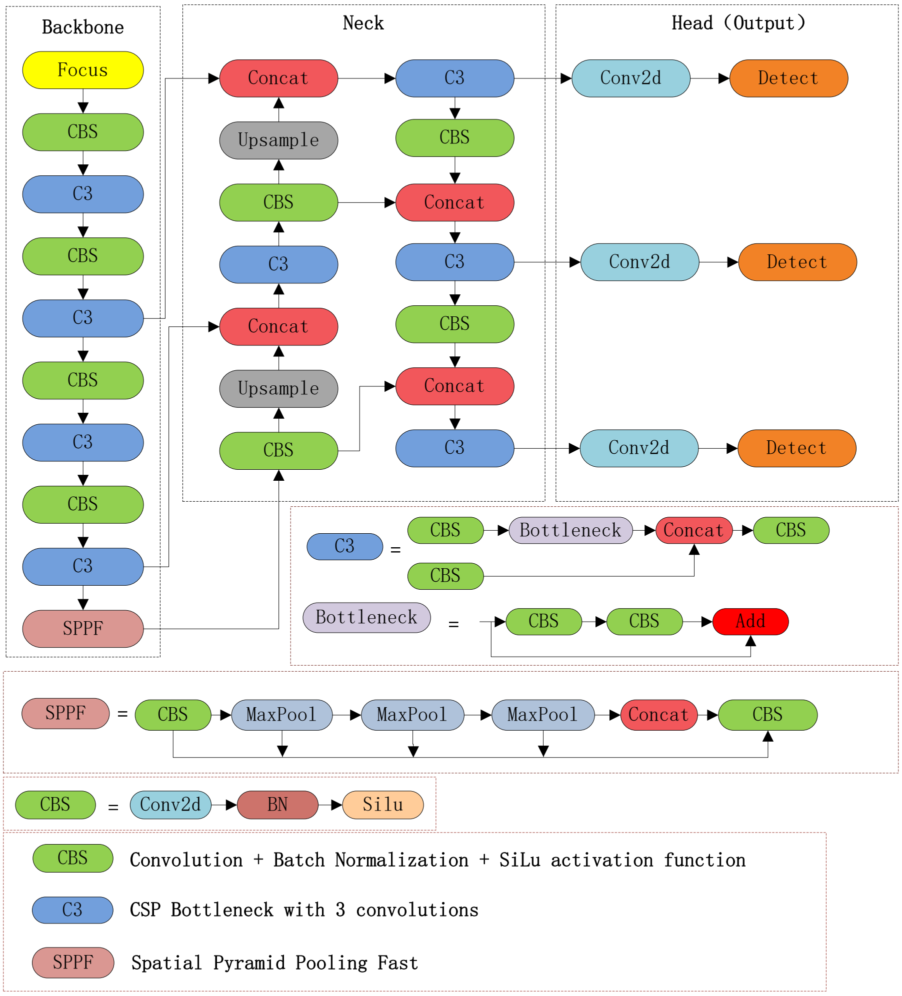


Fig. 2. YOLOv5 network structure

2.1.1 Backbone network

YOLOv5 uses the backbone network of Cross Stage Partial Network (CSP) network structure. CSP network structure includes convolution operation and CSP. The CSP structure uses the idea of DenseNet network structure to combine gradient information into the feature map. This structure can not only effectively use a large number of image features, but also reduce the redundancy in the network structure.

2.1.2 Neck network

YOLOv5 uses the Path Aggregation Network (PAN) + Feature Pyramid Network (FPN) structure in the Neck network part. The FPN structure outputs the feature maps between different layers in the backbone network to obtain the features between different layers. The PAN structure is based on the acquisition of the lower sampling features of the FPN and then adds the upper sampling for feature fusion. In YOLOv5, the PAN+FPN structure can make full use of and fuse the features extracted from the backbone network, so as to obtain better detection performance.

2.1.3 Head network

The head part is usually designed to detect the position and category of the target through the feature map extracted from the backbone network or fused from the neck network. YOLOv5 uses three 1×1 convolution layer replaces the full connection layer for prediction and classification, respectively at 20×20, 40×40 and 80×80. The corresponding category probability, target confidence and prediction frame coordinates of the three anchor boxes are predicted on the feature map of three scales.

2.2 Proposed Method

To improve the detection accuracy of YOLOv5 for solder joint defects, this paper proposes an improved YOLOv5 namely CSSD-YOLO that integrated the newly design CSSD module in the YOLOv5 framework. It is used to improve the Backbone and Neck network of YOLOv5, reduce the loss of local feature information during feature extraction, and reduce the size of model parameters. The CA attention mechanism module is embedded to enhance the precise positioning ability of the backbone network and reduce the missed detection rate of the defective target. K-means++ is used to optimize anchor box sizes and overcome the poor clustering effect and stability caused by the random selection of initial values in K-means algorithm.

2.2.1 Proposed CSSD module

The CSSD module is composed of CS and SD sub-modules, as shown in the Figure 3. The function of the CS (Cascade and Shuffle) module is to reduce the loss of information during feature extraction and to inhibit the growth of model parameters. In some previous work, there are many lightweight networks, such as Mobile-Nets [25] and Xception [26]. One of the important ideas to reduce model parameters is deep separable convolution. Deep separable convolution consists of deepwise convolution and pointwise convolution. Deepwise convolution is used to reduce the amount of calculation and improve the calculation speed. However, due to its independent convolution operation on each channel of the input layer, the information of the feature map in the spatial position is difficult to be effectively utilized, so pointwise convolution is used to weight and combine the feature map in the depth direction to generate a new feature map.

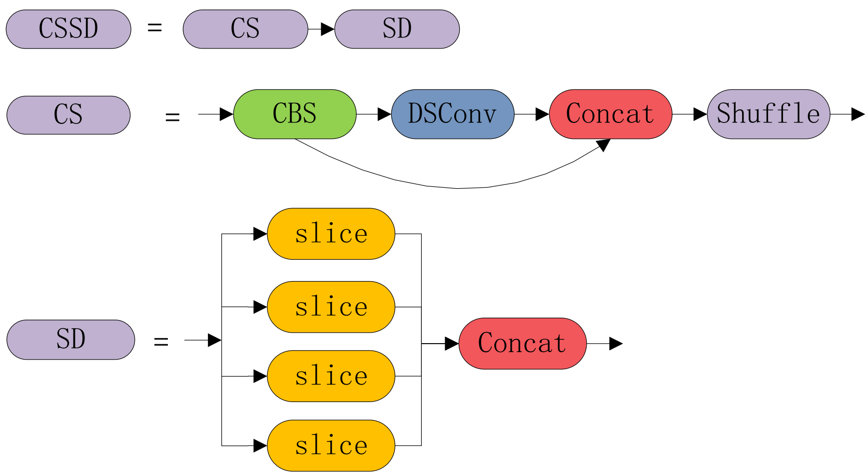


Fig. 3. CSSD module structure

The structure of the CS module is shown in Figure 4. First, the input feature map is convolved (CBS), and then the output feature map is deeply separable convolved (DSConv). Then, the processed feature map is spliced with the output after the first convolution (Concat). Finally, the final output is obtained by shuffling and recombining the spliced results (Shuffle). The purpose of this design is to fuse the features processed by ordinary convolution with the deep separable convolution features, reduce the possibility of local feature information loss, and improve the information association between channels.

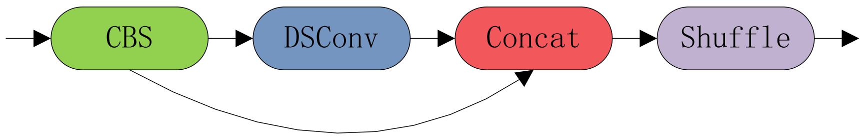


Fig. 4. Composition of the CS block

In the original YOLOv5 framework, the step size of the convolution layer in the CBS module is set to 2. Therefore, the effect after processing is equivalent to the down-sampling operation, which will inevitably lead to the omission of feature information, as shown in Figure 5 (b), especially in the scene with a small area target detection or low resolution. To reduce the information omission and enhance the convolution layer's feature extraction, we set the step size to 1. The process is illustrated in Figure 5 (a). However, this setting is bound to lead to a surge in parameters. Thus, the concept of Space to Depth (SD) is introduced [23]. SD transfers spatial information to a channel, as shown in Eq. (1), thus, avoiding any information loss in the spatial dimension reduction. In the SD process, firstly, each of the other pixel in the feature map input is went through a sampling process. In this way, four groups of feature maps can be obtained as shown in Figure 5 (c). A feature map has 3-dimensions, width (w), height (h) and channel (c). The purpose of SD is to convert the information from w and h to c that is also known as depth (d). When the four groups of feature maps are spliced in the channel direction, the information on the w-h plane is converted to the c dimension, as shown in Figure 5 (c). Such processing can reduce the size of image space without losing feature information.

$$M_s : [0,1]^{sH \times sW \times C} \rightarrow [0,1]^{H \times W \times s^2 C} \quad (1)$$

where,
Ms = Feature map
sH = The height of the input feature map
sW = The width of the input feature map
C = The channel of the input feature map
H = The height of the output feature map
W = The width of the output feature map
S = Pooling block size

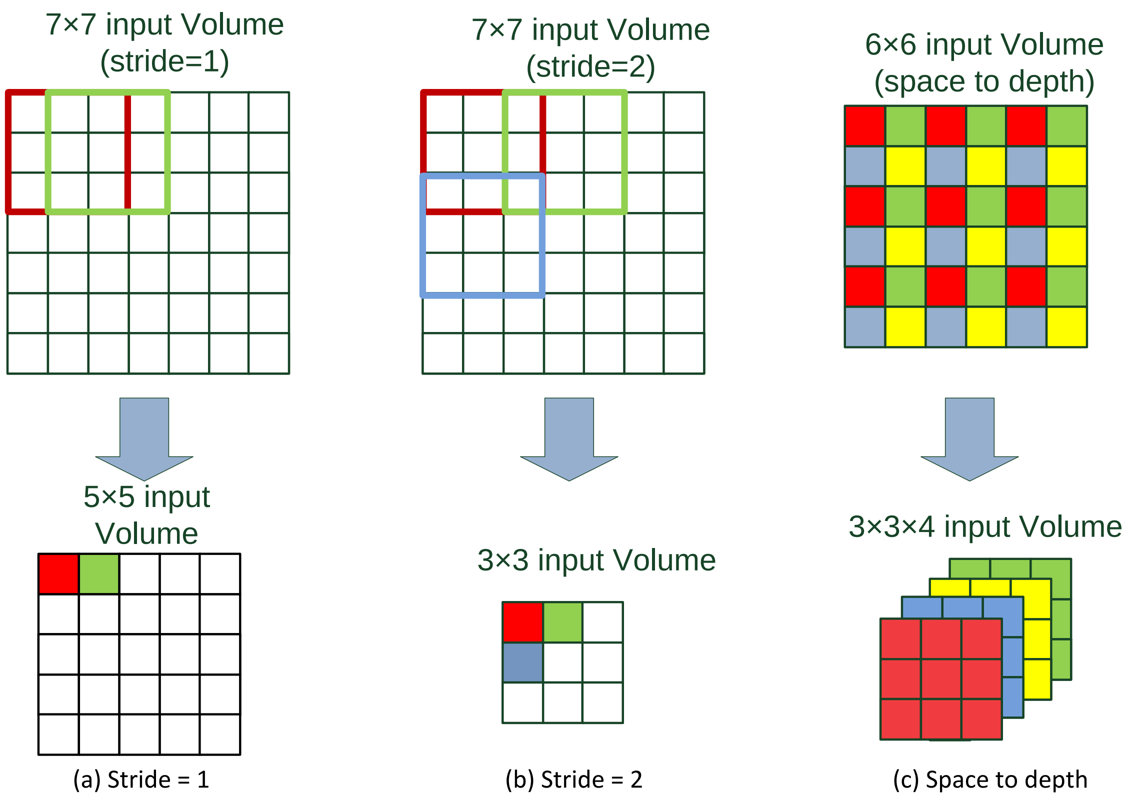


Fig. 5. Three different convolution methods

The SD module is shown in Figure 6 showing that the feature map is divided into four slices before concatenation.

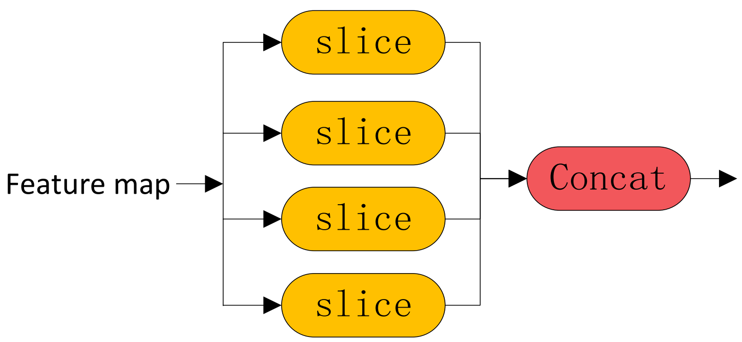


Fig. 6. Composition of the SD block

The integration method of the CSSD is to replace the original CBS module located prior to each C3 module with the CSSD module in the backbone network of the YOLOv5 framework. In the Neck network, the first two CBS modules are replaced with CS sub-modules, and the last two CBS modules are replaced with CSSD modules, as shown in Figure 7.

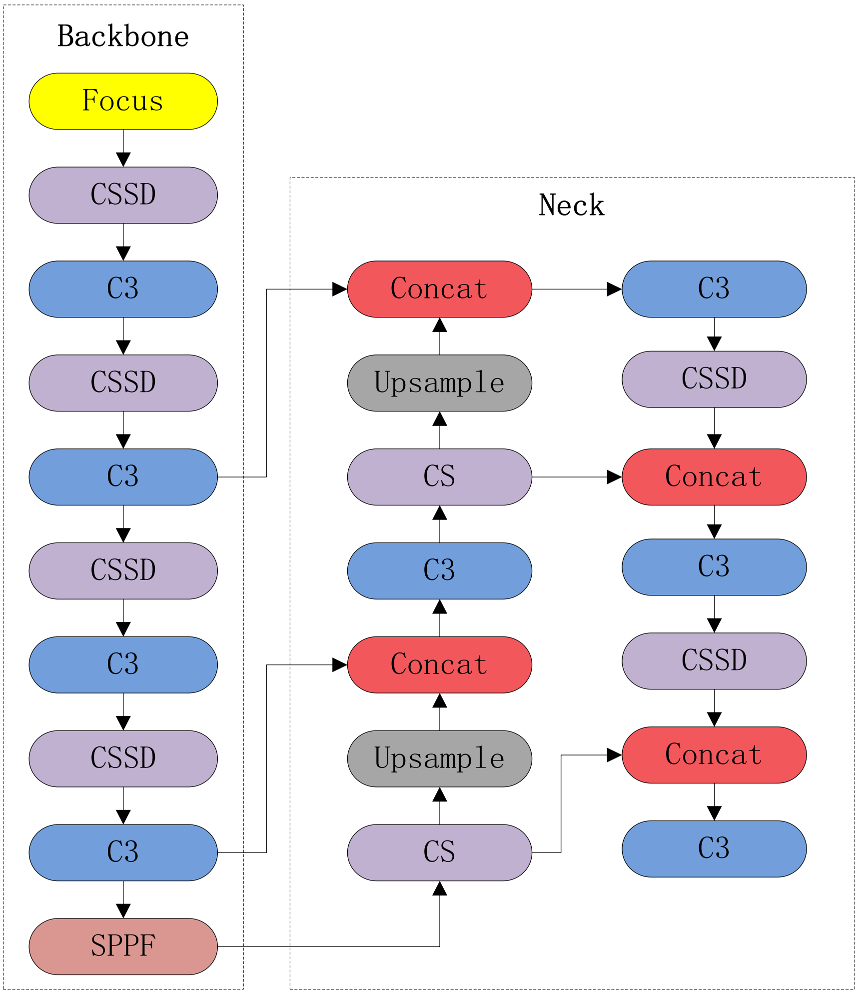


Fig 7. Integration mode of CSSD module

2.2.2 Embedded design of Coordinate Attention attention module

The Coordinate Attention (CA) mechanism [27] has made targeted improvements based on the Squeeze Excitation (SE) attention mechanism [28] and retains the feature location information of the image. Compared with the SE attention module, this module can not only obtain the long-range dependence in the spatial direction but also enhance the position information expression of features and increase the global receptive field of the network. As shown in Figure 8, the input features are pooled by one-dimensional adaptive averaging in the X-axis and Y-axis directions to obtain independent directional sensing features that retain the X-axis and Y-axis information.

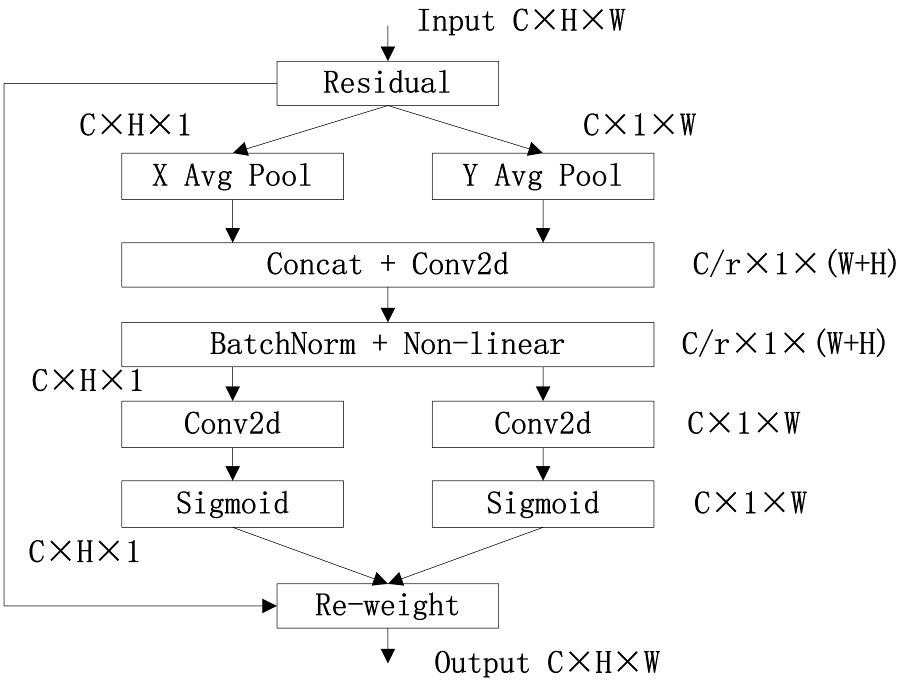
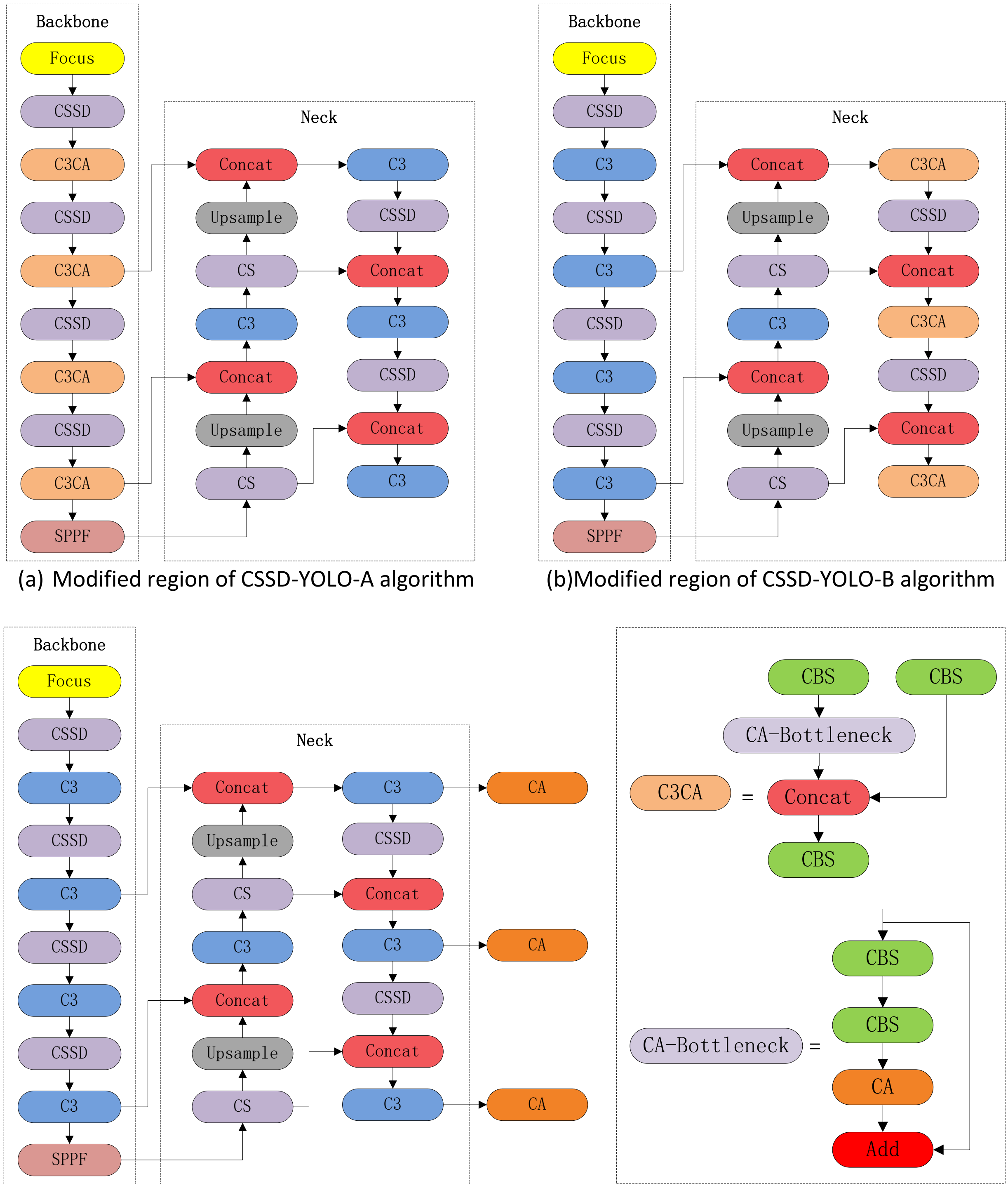


Fig. 8. Coordinate attention block architecture

One spatial direction captures long-range dependence, while the other retains accurate position information. The two one-dimensional features obtained are spliced on the w dimension through a convolution and nonlinear activation function. Then the features are split in the channel dimension. Two feature maps with specific spatial directions and long-range dependence are obtained through the convolution and sigmoid activation functions. These two feature maps can complementarily be applied to the input feature map to enhance the interested target. The feature map with attention weight in the w and h directions is finally obtained through feature fusion with the original feature.

In this paper, three kinds of integration structures of the CA module are designed, and the CA module is embedded in different parts of the improved YOLOv5 structure. These three structures are CSSD-YOLO-A, CSSD-YOLO-B, and CSSD-YOLO-C, as shown in Figure 9. In Figure 9(a), the CA module is embedded in the C3 module to form the C3CA module, as shown in Figure 9(d). Furthermore, all C3 module in the Backbone network is replaced by the C3CA module. In Figure 9(b), replace three output C3 modules in the Neck network with C3CA modules. In Figure 9(c), each output C3 module in the Neck network is directly connected to the CA module.



(a) Modified region of CSSD-YOLO-A algorithm (b)Modified region of CSSD-YOLO-B algorithm

(c) Modified region of CSSD-YOLO-C algorithm (d) C3CA structure

Fig. 9. Three kinds of CSSD-YOLO models embedded in CA modules

2.2.3 Anchors sizes optimization by k-means++

The YOLOv5 algorithm must preset the anchor box sizes before training and prediction. The detection accuracy is affected by the initial anchor setting. It is imperative to preset the anchor in the algorithm. In the original YOLOv5 algorithm, the clustering anchor uses the K-means algorithm. The K-means clustering [29] adaptively sets the anchor size in line with the data set, making the model training easier to converge. However, the results of the K-means algorithm are greatly affected by the selection of initial points. Usually, it requires multiple clustering to obtain more stable convergence results. To obtain a priori frame that is more consistent with the solder joint defect and improve the efficiency of network training, we use the K-means++ algorithm [30] to analyze the solder joint defect dataset and obtain nine groups of anchor box sizes, as shown in Table 1. K-means++ overcomes the poor clustering effect and stability caused by the random selection of initial values in the K-means algorithm. It reduces the error caused by the size of the last frame.

Table 1
Dimensions of anchor box

Feature map	20×20	40×40	80×80
	(34,41)	(21,99)	(55,57)
Dimensions of anchor box	(29,59)	(37,68)	(20,185)
	(37,53)	(26,117)	(32,171)

3.Results and Discussion

3.1 Experimental Framework

The performance of CSSD-YOLO is evaluated on the solder joint dataset. The experiment shows that the design of CSSD-YOLO is reasonable and effective and has practical application value in industrial scenarios.

3.1.1 Dataset

The sample image data set are collected from an electronic storage device production factory. 473 defective SMT chip images are collected by a CCD industrial camera. These SMT chip images include 3154 defective solder joints. There are two kinds of defects, insufficient defect and pin shifting defect.

3.1.2 Experimental environment

The whole system experiments are carried out with Intel(R) Core™ i5-12400F CPU @ 2.50GHz, 16 GB RAM, NVIDIA RTX 3060 12GB display memory and Windows 10 Pro as the operating system. The software environment is CUDA 11.3, cudnn 8.0, Python 3.9, and PyTorch 1.13.1.

3.1.3 Evaluation criterion

At present, the mainstream general indicators for evaluating the performance of object detection algorithms include precision, recall, mAP (Mean Average Precision), parameters (the number of parameters in model), and so on. In this experiment, we choose precision, recall, mAP, and parameters, to evaluate the algorithm. Precision, recall, and mAP are calculated using Eq. (2) to Eq. (4). TP, FP, TN, and FN stand for true positive, false positive, true negative, and false negative.

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$mAP = \frac{\sum_{i=1}^{N} AP_i}{N} AP = \int_0^1 P(R)dR \tag{4}$$

3.2 Experimental Results
3.2.1 Comparative experiment

In order to compare the detection effect of three schemes of attention mechanism integration, a comparative experiment was conducted. In this experiment, the train set and validation set for the solder joint dataset in CSSD-YOLO-A, CSSD-YOLO-B, and CSSD-YOLO-C are used in the three model schemes proposed in Section 2.2.2. The performance of the model is compared and evaluated on the test set. The experimental results are shown in Table 2. On the precision indicator, CSSD-YOLO-B is 1.0% and 2.8% higher than CSSD-YOLO-A and CSSD-YOLO-C, respectively. On the recall indicator, CSSD-YOLO-B is 4.9% and 4.4% higher than CSSD-YOLO-A and CSSD-YOLO-C, respectively. On the mAP index, CSSD-YOLO-B is 3.2% and 1.3% higher than CSSD-YOLO-A and CSSD-YOLO-C, respectively. By comparing the performance of CSSD-YOLO-A, CSSD-YOLO-B, and CSSD-YOLO-C model schemes, CSSD-YOLO-B has the best detection effect. So, we use CSSD-YOLO-B as our final improvement scheme.

Table 2
The experimental results were compared by different models embedded in CA modules

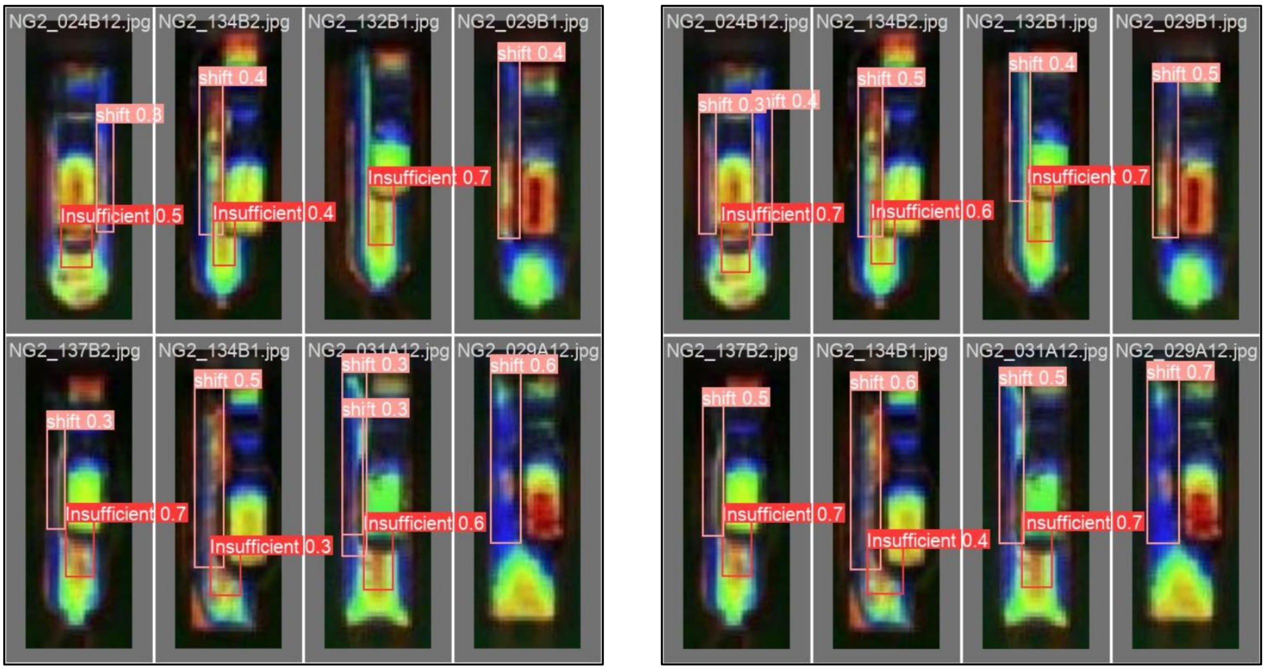
Network model	Precision/%			Recall/%			mAP/%		
	Insufficient	Shifting	All	Insufficient	Shifting	all	Insufficient	Shifting	all
CSSD-YOLO-A	92.5	92.7	92.6	94.5	87.0	90.8	96.5	89.5	93.0
CSSD-YOLO-B	91.6	95.6	93.6	99.5	91.9	95.7	97.2	95.2	96.2
CSSD-YOLO-C	86.3	95.2	90.8	96.6	86.0	91.3	96.0	93.8	94.9

In order to verify the advantages of the improved algorithm, we used the same dataset to train under different networks and used the classic two-stage network Faster R-CNN [31], one-stage network YOLOv4 [32], YOLOv5 [33], YOLOv7 [34] and YOLOv8 [35] for comparative experiments. The experimental results are shown in Table 3. The CSSD-YOLO proposed in this paper performs the best in the three indicators of precision, recall, and mAP among the above six algorithms. Precision has increased 17.4%, 12.2%, 21.9%, 15.4% and 8.9% respectively compared with YOLOv4-tiny, YOLOv5, YOLOv7-tiny, YOLOv8s and Faster R-CNN. Recall increased by 22.1%, 9.4%, 6.6%, 16.0% and 0.6% respectively compared with YOLOv4-tiny, YOLOv5, YOLOv7-tiny, YOLOv8s and Faster R-CNN. Compared with YOLOv4-tiny, YOLOv5, YOLOv7-tiny, YOLOv8s and Faster R-CNN, mAP increased by 14.8%, 9.0%, 14.5%, 12.0% and 13.5% respectively. CSSD-YOLO has the least parameter quantity compared to the above five algorithms, only 5.7M, 1.3M less than the original YOLOv5. Based on the above experimental results, the CSSD-YOLO algorithm proposed in this paper is a solder joint defect detection algorithm with better detection accuracy and smaller model parameters. It has excellent comprehensive performance and high application value.

Table 3
Comparison results of detection performance of different algorithms

	Precision/%			Recall/%			mAP/%			Parameters (M)
	Insufficient	Shifting	all	Insufficient	Shifting	all	Insufficient	Shifting	all	
YOLOv4-tiny	82.3	70.1	76.2	87.4	59.8	73.6	90.6	72.2	81.4	6.1
YOLOv5s	87.7	75.0	81.4	98.6	73.9	86.3	96.5	77.9	87.2	7.0
YOLOv7-tiny	87.2	56.2	71.7	99.5	78.3	89.1	95.6	67.9	81.7	6.2
YOLOv8s	87.8	68.5	78.2	98.2	60.9	79.7	96.6	71.7	84.2	11.1
Faster R-CNN	87.3	82.0	84.7	98.7	91.5	95.1	90.1	75.3	82.7	41.1
CSSD-YOLOv5	91.6	95.6	93.6	99.5	91.9	95.7	97.2	95.2	96.2	5.7

Figure 10 shows the defect detection effect of the two algorithms, YOLOv5s and CSSD-YOLO. The YOLOv5s algorithm has missed detection and does not detect the pin shifting defect of the first solder joint and the third solder joint, while the CSSD-YOLO algorithm has good detection effect and can meet the accuracy rate of solder joint defect detection in actual production.



(a) YOLOv5s (b) CSSD-YOLO

Fig. 10. Comparison of the detection results of YOLOv5s (a) and CSSD-YOLO(b) in the solder joint dataset

3.2.2 Ablation tudy

In order to more comprehensively analyze the advantages of each improved module in CSSD-YOLO for solder joint defect detection, this paper designs the ablation experiment based on the original YOLOv5. The K-means++, CSSD module, and CA attention mechanism module are added to the original YOLOv5, respectively, and the original algorithm is used as the control group. The specific experimental content and test results are shown in Table 4. Using the K-means++ algorithm to select priori anchor boxes will improve the detection precision by 3.8%, recall by 6%, and mAP by 6%. Integrating the CSSD module into the Backbone and Neck network reduces the loss of feature information, reduces the number of parameters, and improves the detection effect of small defects and low-resolution defects compared with the original YOLOv5 algorithm. Increase detection precision by 8.9%, recall by 6.9%, and mAP by 6%. The CA attention mechanism module is integrated into the C3 module in the Neck network to improve the network representation ability while highlighting the characteristics of critical locations and effectively solving the problem of harrowing feature extraction. Compared with the original YOLOv5, the detection precision is improved by 6.4%, recall by 4.6%, and mAP by 7.2%. Through the comparison of ablation experiments, it is found that the performance improvement of the model after adding each improved module, CSSD-YOLO proposed in this paper, is the most significant. Both types of defect detection objects have achieved good detection results, which verifies the effectiveness of the algorithm in this paper for solder joint defect recognition.

Table 4
Ablation study on solder joint dataset

	Precision/%			Recall/%			mAP/%		
	Insufficient	Shifting	all	Insufficient	Shifting	all	Insufficient	Shifting	all
YOLOv5	87.7	75.0	81.4	98.6	73.9	86.3	96.5	77.9	87.2
YOLOv5 + Kmeans++	90.3	80.2	85.2	98.6	88.0	93.3	97.2	89.2	93.2
YOLOv5 + CSSD	91.4	89.2	90.3	99.5	87.0	93.2	98.4	88.0	93.2
YOLOv5 + CA	91.2	84.4	87.8	99.2	82.6	90.9	98.9	89.9	94.4
YOLOv5 + CSSD + CA+ Kmeans++	91.6	95.6	93.6	99.5	91.9	95.7	97.2	95.2	96.2

4. Conclusions

This research has developed a high-precision detection of two common defects in the SMT solder joint defect dataset, which utilizes a modified deep-learning algorithm of YOLOv5. A CSSD module is proposed, integrated into the backbone and neck networks of YOLOv5 architecture to improve the feature extraction ability and reduce the model parameters. The CA attention mechanism module's embedded structure is designed to improve further the feature extraction ability and positioning accuracy of the network for solder joint defects. K-means++ algorithm is used to cluster the prior anchor boxes instead of the K-means algorithm, and the best anchor box sizes are obtained, which improves the detection accuracy. Finally, the experimental results show that the improved method proposed in this paper can accurately detect different defects in the solder joint image. Compared with the original YOLOv5s under the same test conditions, the method in this paper improves the precision by 12.2%, recall by 9.4%, mAP by 9.0%, and model parameters reduced by 1.3M. Compared with the general object detection model, it also significantly improved. This method provides valuable help for solder joint defect detection. However, there are still some things that could be improved. The model proposed in this paper still has room for improvement in the detection precision of defects. The detection ability for small defect targets needs to be further enhanced. Future research will continue to improve the detection effect of the model on small targets and further improve the detection accuracy of solder joint defects.

Acknowledgment
This research was funded by a grant from Universiti Teknologi MARA Cawangan Melaka (TEJA Grant 2022 GDT 2023/1-14).

References
[1] Nordin, Nur Fatihah, Kee Quen Lee, and Hooi Siang Kang. "Energy Harvesting of Daily Human Life Activities using a Self-Made Piezoelectric System." *Progress in Energy and Environment* 10 (2019): 1-5.

[2] Wong, F., E. H. Loh, P. Y. Lim, R. R. Porle, R. Chin, K. Teo, and K. A. Mohamad. "Development of an Electronic Glove with Voice Output for Finger Posture Recognition." *Journal of Advanced Research in Applied Mechanics* 10: 36-46.

[3] Dai, Wenting, Abdul Mujeeb, Marius Erdt, and Alexei Sourin. "Soldering defect detection in automatic optical inspection." *Advanced Engineering Informatics* 43 (2020): 101004. https://doi.org/10.1016/j.aei.2019.101004

[4] Zhang, Huan, Liangxiao Jiang, and Chaoqun Li. "CS-ResNet: Cost-sensitive residual convolutional neural network for PCB cosmetic defect detection." *Expert Systems with Applications* 185 (2021): 115673. https://doi.org/10.1016/j.eswa.2021.115673

[5] Cai, Li, and Jingchuan Li. "PCB defect detection system based on image processing." In *Journal of Physics: Conference Series*, vol. 2383, no. 1, p. 012077. IOP Publishing, 2022. https://doi.org/10.1088/1742-6596/2383/1/012077

[6] Liu, Haiying, Fengqian Sun, Jason Gu, and Lixia Deng. "Sf-yolov5: A lightweight small object detection algorithm based on improved feature fusion mode." *Sensors* 22, no. 15 (2022): 5817. https://doi.org/10.3390/s22155817

[7] Wu, Hao, Xianmin Zhang, Hongwei Xie, Yongcong Kuang, and Gaofei Ouyang. "Classification of solder joint using feature selection based on Bayes and support vector machine." *IEEE Transactions on Components, Packaging and Manufacturing Technology* 3, no. 3 (2013): 516-522. https://doi.org/10.1109/TCPMT.2012.2231902
[8] Xiao, Meng, Nian Cai, Zhuokun Mo, Shule Yan, Nili Tian, Jing Ma, and Han Wang. "IC solder joint inspection via adaptive statistical modeling." *Soldering & Surface Mount Technology* 35, no. 3 (2023): 134-142. https://doi.org/10.1108/SSMT-12-2021-0069
[9] Dai, Wenting, Abdul Mujeeb, Marius Erdt, and Alexei Sourin. "Towards automatic optical inspection of soldering defects." In *2018 International Conference on Cyberworlds (CW)*, pp. 375-382. IEEE, 2018. https://doi.org/10.1109/CW.2018.00074
[10] Zhang, Kaihua, and Haikuo Shen. "Solder joint defect detection in the connectors using improved faster-rcnn algorithm." *Applied Sciences* 11, no. 2 (2021): 576. https://doi.org/10.3390/app11020576
[11] Ding, Runwei, Linhui Dai, Guangpeng Li, and Hong Liu. "TDD-net: a tiny defect detection network for printed circuit boards." *CAAI Transactions on Intelligence Technology* 4, no. 2 (2019): 110-116. https://doi.org/10.1049/trit.2019.0019
[12] Han, H., Y. Xu, B. Sun, C. He, and G. Liao. "Using active thermography for defect detection of aerospace electronic solder joint base on the improved Tiny-YOLOv 3 network." *Chin. J. Sci. Instrum.* 41, no. 11 (2020): 42-49.
[13] Marquez, Enrique S., Jonathon S. Hare, and Mahesan Niranjan. "Deep cascade learning." *IEEE transactions on neural networks and learning systems* 29, no. 11 (2018): 5475-5485. https://doi.org/10.1109/TNNLS.2018.2805098
[14] Cai, Nian, Guandong Cen, Jixiu Wu, Feiyang Li, Han Wang, and Xindu Chen. "SMT solder joint inspection via a novel cascaded convolutional neural network." *IEEE Transactions on Components, Packaging and Manufacturing Technology* 8, no. 4 (2018): 670-677. https://doi.org/10.1109/TCPMT.2018.2789453
[15] Wang, Jing, and Qingwei Zhang. "Visual Defect Detection for Substation Equipment based on Joint Inspection Data of Camera and Robot." In *2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pp. 491-495. IEEE, 2020. https://doi.org/10.1109/ITOEC49072.2020.9141563
[16] Ye, Juan, Linyan Wang, Yaqi Wang, Lirong Chen, Yuan Huang, Fengbo Huang, Davin C. Ashraf Ashraf *et al.,* "A Deep Learning Approach with Cascade-network Design for Eyelid Tumors Diagnosis Based on Gigapixel Histopathology Images." (2022). https://doi.org/10.21203/rs.3.rs-1597378/v1
[17] Tang, Yi, Wenbin Zou, Zhi Jin, Yuhuan Chen, Yang Hua, and Xia Li. "Weakly supervised salient object detection with spatiotemporal cascade neural networks." *IEEE Transactions on Circuits and Systems for Video Technology* 29, no. 7 (2018): 1973-1984. https://doi.org/10.1109/TCSVT.2018.2859773
[18] Tianyang, Dong, Zhang Jian, Gao Sibin, Shen Ying, and Fan Jing. "Single-tree detection in high-resolution remote-sensing images based on a cascade neural network." *ISPRS International Journal of Geo-Information* 7, no. 9 (2018): 367. https://doi.org/10.3390/ijgi7090367
[19] Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788. 2016. https://doi.org/10.1109/CVPR.2016.91
[20] Yao, Jia, Jiaming Qi, Jie Zhang, Hongmin Shao, Jia Yang, and Xin Li. "A real-time detection algorithm for Kiwifruit defects based on YOLOv5." *Electronics* 10, no. 14 (2021): 1711. https://doi.org/10.3390/electronics10141711
[21] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251-1258. 2017. https://doi.org/10.1109/CVPR.2017.195
[22] Ma, Ningning, Xiangyu Zhang, Hai-Tao Zheng, and Jian Sun. "Shufflenet v2: Practical guidelines for efficient cnn architecture design." In *Proceedings of the European conference on computer vision (ECCV)*, pp. 116-131. 2018. https://doi.org/10.1007/978-3-030-01264-9_8
[23] Sajjadi, Mehdi SM, Raviteja Vemulapalli, and Matthew Brown. "Frame-recurrent video super-resolution." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 6626-6634. 2018. https://doi.org/10.1109/CVPR.2018.00693
[24] Yuan, Jianfei, and Yongkang Peng. "Defect Detection Method of PCB Based on Improved YOLOv5."
[25] Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).
[26] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251-1258. 2017. https://doi.org/10.1109/CVPR.2017.195
[27] Hou, Qibin, Daquan Zhou, and Jiashi Feng. "Coordinate attention for efficient mobile network design." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 13713-13722. 2021. https://doi.org/10.1109/CVPR46437.2021.01350

[28] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132-7141. 2018. https://doi.org/10.1109/CVPR.2018.00745
[29] Likas, Aristidis, Nikos Vlassis, and Jakob J. Verbeek. "The global k-means clustering algorithm." *Pattern recognition* 36, no. 2 (2003): 451-461. https://doi.org/10.1016/S0031-3203(02)00060-2
[30] Arthur, David, and Sergei Vassilvitskii. "K-means++ the advantages of careful seeding." In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pp. 1027-1035. 2007.
[31] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems* 28 (2015).
[32] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: Optimal speed and accuracy of object detection." *arXiv preprint arXiv:2004.10934* (2020).
[33] Jocher, Glenn, Ayush Chaurasia, Alex Stoken, Jirka Borovec, Yonghye Kwon, Jiacong Fang, Kalen Michael *et al.,* "ultralytics/yolov5: v6. 1-tensorrt, tensorflow edge tpu and openvino export and inference." *Zenodo* (2022).
[34] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." *arXiv e-prints* (2022): arXiv-2207.
[35] Jocher, Glenn, Alex Stoken, Jirka Borovec, Liu Changyu, Adam Hogan, Laurentiu Diaconu, Jake Poznanski *et al.,* "ultralytics/yolov5: v3. 0." *Zenodo* (2020).