



ARTICLE

Printed Circuit Board (PCB) Surface Micro Defect Detection Model Based on Residual Network with Novel Attention Mechanism

Xinyu Hu, Defeng Kong*, Xiyang Liu, Junwei Zhang and Daode Zhang

School of Mechanical Engineering, Hubei University of Technology, Wuhan, 430068, China

*Corresponding Author: Defeng Kong. Email: kongdefeng0916@gmail.com

Received: 28 September 2023 Accepted: 23 November 2023 Published: 30 January 2024

ABSTRACT

Printed Circuit Board (PCB) surface tiny defect detection is a difficult task in the integrated circuit industry, especially since the detection of tiny defects on PCB boards with large-size complex circuits has become one of the bottlenecks. To improve the performance of PCB surface tiny defects detection, a PCB tiny defects detection model based on an improved attention residual network (YOLOX-AttResNet) is proposed. First, the unsupervised clustering performance of the K-means algorithm is exploited to optimize the channel weights for subsequent operations by feeding the feature mapping into the SENet (Squeeze and Excitation Network) attention network; then the improved K-means-SENet network is fused with the directly mapped edges of the traditional ResNet network to form an augmented residual network (AttResNet); and finally, the AttResNet module is substituted for the traditional ResNet structure in the backbone feature extraction network of mainstream excellent detection models, thus improving the ability to extract small features from the backbone of the target detection network. The results of ablation experiments on a PCB surface defect dataset show that AttResNet is a reliable and efficient module. In Torify the performance of AttResNet for detecting small defects in large-size complex circuit images, a series of comparison experiments are further performed. The results show that the AttResNet module combines well with the five best existing target detection frameworks (YOLOv3, YOLOX, Faster R-CNN, TDD-Net, Cascade R-CNN), and all the combined new models have improved detection accuracy compared to the original model, which suggests that the AttResNet module proposed in this paper can help the detection model to extract target features. Among them, the YOLOX-AttResNet model proposed in this paper performs the best, with the highest accuracy of 98.45% and the detection speed of 36 FPS (Frames Per Second), which meets the accuracy and real-time requirements for the detection of tiny defects on PCB surfaces. This study can provide some new ideas for other real-time online detection tasks of tiny targets with high-resolution images.

KEYWORDS

Neural networks; deep learning; ResNet; small object feature extraction; PCB surface defect detection

1 Introduction

Printed circuit board (PCB) is the basic carrier of electronic components, due to the many complex production processes, harsh production environment, as well as the instability of the production equipment, the production and manufacturing process often appear missing hole, mouse bite, open



circuit, short, spur, spurious copper and other defects [1]. These defects not only affect the performance of electronic products but may even lead to serious failure of electronic products and scrap [2]. Therefore, to improve the quality of PCB boards and production efficiency, the detection of common defects is the basic requirement of PCB manufacturing quality control.

Currently, PCB defect detection is mainly through manual visual inspection combined with microscope observation of defects, or Automated Optical Inspection (AOI) equipment to determine the type and location of defects [3]. The manual visual inspection method is limited by the detection efficiency and accuracy and has been slowly abandoned. As a non-contact surface defects online inspection technology, AOI has been widely used in the PCB manufacturing process, and its core detection algorithm is divided into traditional machine vision detection algorithm and deep learning-based detection algorithm [4]. Traditional machine vision-based inspection algorithms usually require complex artificial feature extraction, and then select a classifier for defect detection, or design a standard template to match the defective features. With the development of electronic products to miniaturization and high-density integration, PCB circuit board wiring has become complex and tiny, and there are problems such as complex production of standard templates, difficulty in manual extraction of defective features, and the need to detect different models of PCBs when There are problems such as complex standard stencil production, difficulty in manual extraction of defective features, and the need to re-model when detecting different PCB models [5–8].

In contrast, deep learning-based target detection algorithms automatically extract features from image features with displacement, scale, and deformation invariance, which can effectively solve the problems of complex template production, difficult feature extraction, and low flexibility of traditional machine vision algorithms [9]. In recent years, many scholars have proposed deep learning-based PCB surface defect detection algorithms for PCB surface defect characteristics, to break through the bottleneck of traditional machine vision defect detection technology. For the problem of imbalance in the distribution of defect types on the PCB surface, and the problem of different misclassification costs of real defects and pseudo defects, Zhang et al. added a cost-sensitive adjustment layer in ResNet, and according to the imbalance degree, assigned larger weights to a few real defects, and optimized the CS-ResNet detection network by minimizing the weighted cross-entropy loss function [8]; Wu et al. introduced MobileNetV3 as a feature extraction network and Inceptionv3 as a detection network, respectively, and then used an improved clustering algorithm to determine the anchor frames, which enhanced the performance of PCB defects detection, and the light-weight transformation for PCB defects with complex background [10]; Yang et al. improved the detection accuracy of defects of different sizes by increasing the detection head, and introduced the convolutional block attention module (CBAM) attention module in order to enhance the image information in complex background [11]; Hu et al. proposed an improved detection model of PCB tiny defects based on the Faster RCNN architecture, and used the optimized ROI-Pooling layer to extract the ROI features, while adopting a dual fully connected layer in the ROI region to classify the regression target, in addition to adding filtering features of the pyramid network to enhance the ability to mimic the multi-scale and irregular defect features and improve the applicability of the detection model [12]; To solve the problems of low detection efficiency, high memory consumption, and low sensitivity to small defects in the detection of surface defects of PCB, a new deep learning-based lightweight detection model is proposed to improve the detection of small defects in the surface defects of PCB. Propose a new lightweight defect detection network based on deep learning, which incorporates the improved CSPDarknet module on the basis of YOLOX and develops the coordinated attention mechanism (YOLOX-MC-CA), YOLOX-MC-CA is developed on the basis of YOLOX, which utilizes the polar coordinate attention (CA) mechanism to improve the PCB small surface defects Recognition ability [13]; Xia et al. proposed a PCB small defect

detection method based on focal loss and Region-based Fully Convolutional Network (FL-RFCN) and Parallel High Definition Feature Extraction (PHFE) for the problem of imbalance between the foreground and the background of PCB small defects, which improves the PCB small defects detection accuracy [14]; Tang et al. constructed a new PCB-YOLO detection algorithm using the YOLO algorithm as the basic framework, which realizes the real-time PCB surface defects detection requirements by incorporating the K-means++ algorithm, the Swin transformer module, the multi-scale detection head, the Efficient-IoU (EIoU) loss function, and the depth separable convolution [15]. The deep learning-based detection method uses different sizes of sensory wild convolution kernels to perform convolution operations on the input image to obtain the high semantic features of the target and then decodes the feature vectors to obtain the classification results and location information. In actual production, PCB tiny defects in the image account for a relatively small percentage, the defect location is random and irregular, especially in large-scale images, PCB tiny defects pixels account for a lower percentage, the detection is more and more difficult, so based on the deep learning model extracted from the tiny defects feature information has a discretization characteristic.

Existing PCB tiny defect detection based on deep learning has the following problems: (1) the deep learning model of the convolutional neural network is prone to lose the key information of the tiny target in multiple convolution operations on the image, the deeper the network, the more convolutions are done, the greater the possibility of losing the key information, resulting in a serious feature loss phenomenon for the small target that originally lacks the expression of the information; (2) the PCB surface tiny defects feature information is small, easily confused with the background, the deep network in the feature extraction process, failed to effectively remove the redundant background information, resulting in the discretization of the tiny defects feature vector, the characterization ability is not strong.

To solve the above problems, this paper proposes an improved AttResNet basic residual module and, based on the YOLOX model, incorporates the AttResNet module into the backbone network to the performance of feature extraction for tiny PCB defects, thus obtaining an improvement in detection accuracy. In this paper, we start with the structure of ResNet, which is the most basic unit of the deep learning network, which is very different from other PCB surface defect detection literature in the past. A large amount of literature focuses on the construction of the overall deep learning model, such as alternating the stacking of the module structure [16,17] or designing the auxiliary modules [18] to achieve the improvement of efficiency and accuracy, and neglects that ResNet, the basic unit of the composition of the deep network, needs to be corrected as well [19]. The current deep learning frameworks based on convolutional neural networks all benefit from the excellent design of the ResNet network, which allows the network to deepen without the gradient disappearing and network degradation leading to failure, but from our research, we found that the ResNet network also has some drawbacks, such as jumping directly from the line of the shallow feature layer to the deeper feature layer of the information without distinction, which contains both The AttResNet module proposed in this paper is to solve the problem of feature discretization caused by redundant information transfer. The main approach is to implant the K-means-SENet [20,21] module at the direct jumping end of the ResNet network to attenuate the propagation of redundant information, to obtain more robust feature extraction. The module enables the model to improve the accuracy of small target detection, which is a new and fundamentally innovative design that has not appeared in the previous literature. The idea of attention is introduced because the attention mechanism has been a great success in deep learning models, and deep models with the attention mechanism tend to have better performance, for example, the representative channel attention network SENet, the spatial attention network and the hybrid attention network CBAM. It is inspired by the human biological visual system, which pays more

attention to the important parts of interest when the human eye ingests the surrounding information, and this mechanism is consistent with the design idea of this paper to eliminate redundant information and refine the important features, so it can be borrowed, and the main contributions of this paper are summarized as follows. The main contributions of this paper are summarized as follows.

(1) Aiming at the problem of the loss of key feature information easily caused by the convolution process of small feature information, in the SENet channel attention network, before the global average pooling, the K-means algorithm is used to do the clustering operation on the feature map to optimize the weight coefficients and to improve the amount of information of the weak features;

(2) Aiming at the problems of little feature information of tiny defects on the PCB surface, similarity with the background, and the difficulty of eliminating redundant information in the process of deep network feature extraction, the AttResNet module is designed;

(3) To verify the good combinability of the AttResNet module, it is fused with the current mainstream excellent target detection model, and screened out the model that meets the accuracy and efficiency of PCB surface online defect detection, i.e., YOLOX-AttResNet.

The rest of this paper is organized as follows: Detailed description of the PCB surface micro-defect detection model based on a novel attention mechanism residual network in [Section 2](#). In [Section 3](#), the AttResNet components are fused into the backbone of the mainstream target detection framework, and the improved target detection network is implemented on the PCB surface defect dataset for ablation study and result analysis. Finally, [Section 4](#) gives the conclusion and an outlook for future work.

2 Materials and Methods

In this section, a K-means clustering algorithm is introduced into the input feature layer of the attention mechanism module first presented. It is worth noting that clustering is applicable in channel and spatial attention mechanism networks. Second, the enhanced residual network AttResNet is proposed, and the AttResNet theoretical derivation is presented in detail. Finally, the AttResNet components are fused in the backbone of the target detection framework to form a new bottleneck structure for network stacking and deepening.

2.1 Improved Attention Mechanism Network K-Means-SENet

The improved K-means-SENet structure of the attention mechanism network is shown in [Fig. 1](#), and the improvement is focused in the dashed box. Before converting the feature map U into M , a K-means clustering is carried out. The clustering in the dashed box helps to resolve the main information of each channel of the feature map and optimize the weight coefficient.

In [Fig. 1](#), K is a new feature map obtained from U after K-means clustering. The K-means algorithm is the most classical division-based clustering method. Combining image regions or pixels into meaningful parts to eliminate redundant information can obtain a better feature representation of the image. The basic idea of K-means clustering on feature map U is as [Eq. \(1\)](#), that is, to minimize the sum of the squared distances between each pixel point in the class and its corresponding cluster center.

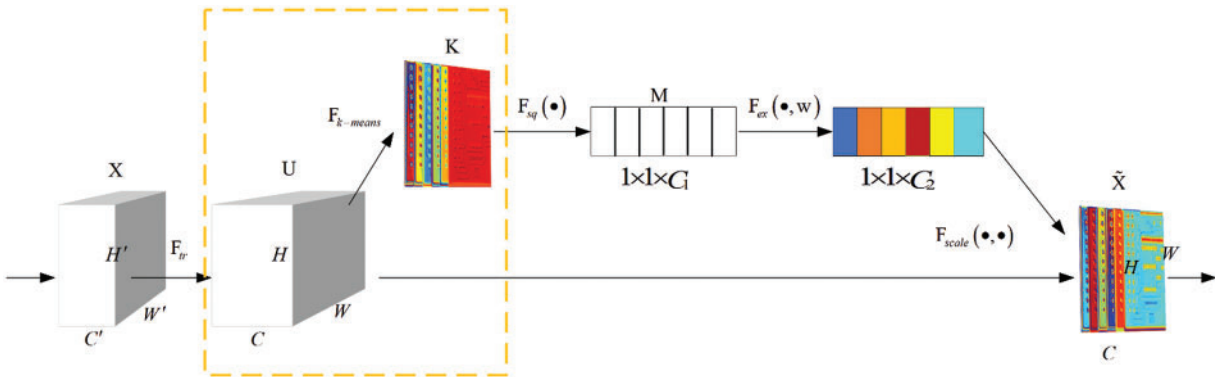


Figure 1: K-means improved the SENet structure diagram

To visualize the clustering effect of K-means on the effective information of feature maps, before and after the introduction of K-means in the channel attention mechanism network (SENet), the feature maps of the 3rd layer of the network (i.e., the feature maps of the ARN_1 layer) were visualized and the differences were compared, and the results of the comparison are shown in Fig. 2. As can be seen from Fig. 2, the main feature information of the feature graph is more obvious after the introduction of K-means.

$$J = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|C(n) - \mu_k\|^2 \tag{1}$$

where n is the coordinate of a pixel in the image; $C(n)$ is the pixel value of the pixel; N is the samples of color data, and K is the number of pixel categories. k is the clustering center of each class of pixels. γ_{nk} is a two-component, which represents whether the current pixel value belongs to the category k : if $\gamma_{nk} = 0$, the current pixel value does not belong to the category k ; if $\gamma_{nk} = 1$, the current pixel value belongs to the first category. To obtain the globally optimal pixel clustering results, more clustering is required for a new feature map K .

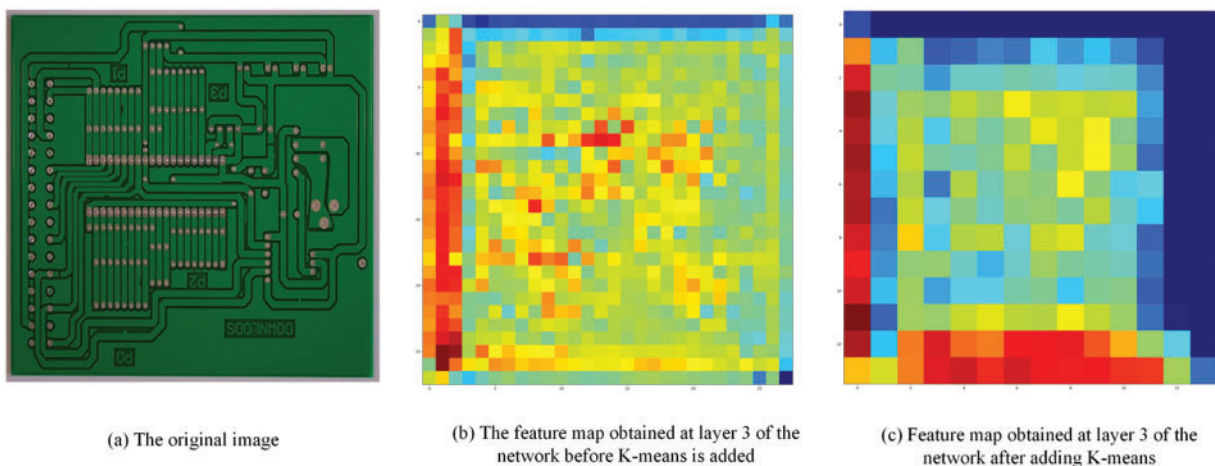


Figure 2: Comparison of network layer 3 feature map (ARN_1 layer) after adding K-means

The feature map K that completes the pixel value update continues with the subsequent global information embedding. The $1 \times 1 \times C1$ feature map M is first obtained by global average pooling of the feature map K . Then the relationship between the channels is learned for the feature map M to obtain the weights $C2$ for different channels. Finally, $C2$ is multiplied by the original feature map U to obtain a new feature map after processing the \tilde{x} . This step can be represented as follows:

$$F_{scale}(u_c, s_c) = s_c \times u_c \quad (2)$$

$$\tilde{x} = F_{scale}(u_c, s_c) \quad (3)$$

where s_c is the learned activation value of each channel; u_c is the original input feature map; \tilde{x} is the new feature map after k-means-SENet processing. The new feature map \tilde{x} has better characterization and is a good bridge for subsequent improvements of the foundation ResNet.

2.2 Enhanced Residual Network AttResNet

The emergence of convolutional neural networks has brought machine vision to a new field, especially with the introduction of the AlexNet network model in 2012. In computer vision, high semantic features need to be obtained by deep network convolution, and the network depth is an important factor in achieving optimal performance. However, as the network model deepens, two other serious problems, gradient disappearance, and network degradation, appear accordingly. The ResNet network was proposed by Kaiming He to solve these problems in deep networks [22], and add shallow redundant information to the higher feature layers for high semantic feature discrete. To this end, an enhanced residual network AttResNet is designed by combining the K-means-SENet network, and the design method is described below.

The residual network is composed of a series of underlying residual blocks. It can be expressed by Eq. (4):

$$x_{l+1} = x_l + F(x_l, W_l) \quad (4)$$

The residual block is divided into two parts: the direct mapping part and the residual part. x_l is directly, $F(x_l, W_l)$ is the residual part, which is usually composed of two or three convolution operations, i.e., the part with the convolution on the upper side of Fig. 2. $F(\cdot)$ is the activation function with ReLU.

For a deeper layer L , its relation to layer l can be expressed as Eq. (5):

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, W_i) \quad (5)$$

The nature of the residual network can be seen from Eq. (5), i.e., the layer x_L can be expressed as the sum of any layer x_l that is shallower than it and the residual part between them. According to the deviation chain rule used in backpropagation, the gradient of the loss function ε concerning x_l can be expressed as:

$$\begin{aligned}
 \frac{\partial \varepsilon}{\partial x_l} &= \frac{\partial \varepsilon}{\partial x_L} \frac{\partial x_L}{\partial x_l} \\
 &= \frac{\partial \varepsilon}{\partial x_L} \left(1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i) \right) \\
 &= \frac{\partial \varepsilon}{\partial x_L} + \frac{\partial \varepsilon}{\partial x_L} \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i)
 \end{aligned} \tag{6}$$

As shown in Eq. (6), 1) $\frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i)$ cannot be -1 all the time in the training process, i.e., there is no gradient disappearance in the residual network; 2) the gradient of a layer x_L can be directly passed to any layer x_l that is shallower than it.

According to the two processes, forward and reverse, of residual networks, the information can be transmitted to each other smoothly between higher and lower levels. In this way, the residual network can solve the problems of gradient disappearance and network degradation in deep networks. It can be seen from Eq. 6 that in the training, the gradient of the shallow layer is calculated by adding a fixed value of 1 to the term $\frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i)$ and taking the partial derivative of x_l . This results in another problem the redundant information in layer x_l is passed to layer x_{l+1} indiscriminately, which is not conducive to extracting high semantic information from the deep layers of the network, affecting the convergence speed of network training. To solve this problem, an adaptive mapping residual network AttResNet is proposed, and its structure is shown in Fig. 3. AttResNet network retains the residual part of the original ResNet, as shown in the right branch of Fig. 3. The K-means-SENet module is inserted in the direct mapping part, as shown in the left branch of Fig. 3.

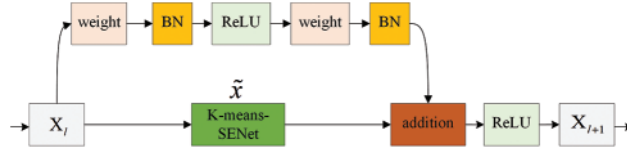


Figure 3: K-means-SENet improved ResNet structure

In Fig. 3, the \tilde{X} is the new feature map in Eq. (3) after K-means-SENet processing. Adding \tilde{X} to the direct mapping edge not only solves the problem of gradient disappearance in the back-propagation process of the network but also adaptively filters out the redundant information in the convolution process and retains the important feature information in the forward propagation process. Eq. (7) is obtained from the improved attention mechanism network K-means-SENet Eq. (3) and the residual network Eq. (4):

$$x_{l+1} = \tilde{x}_l + F(x_l, W_l) \tag{7}$$

Similarly, the relationship between the deep and shallow layers of the residual network Eq. (8) is obtained according to Eq. (5):

$$x_L = \tilde{x}_l + \sum_{i=l}^{L-1} F(\tilde{x}_i, W_i) \tag{8}$$

Substituting Eq. (3) into Eq. (8) to obtain Eq. (9):

$$x_L = F_{scale}(u_c, s_c) + \sum_{i=l}^{L-1} F(\tilde{x}_i, W_i) \quad (9)$$

Eq. (10) is obtained from Eq. (2):

$$x_L = s_c \times x_l + \sum_{i=l}^{L-1} F(\tilde{x}_i, W_i) \quad (10)$$

According to the derivative chain rule used in backpropagation, the gradient of the loss function z concerning x_l can be expressed as in Eq. (11):

$$\frac{\partial \varepsilon}{\partial x_l} = \frac{\partial \varepsilon}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial \varepsilon}{\partial x_L} \left(s_c + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i) \right) \quad (11)$$

where s_c is the learned coefficient of each channel weight of the feature layer, and $s_c \neq 0$. This set of values can adaptively change the corresponding coefficients with different feature layers. It can be analyzed from Eq. (11) that the gradient of the loss function ε concerning x_l is re-skewed from general $\frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, W_i) + 1$ to x_l . The corrected gradient formula will not be +1 directly but will be summed to an adaptive variable value s_c value. This indicates that the AttResNet network learns useful information autonomously in training. This will allow the network to pay more attention to important features, ignore the learning of redundant information, hinder the propagation of unimportant information, accelerate network training, and improve network robustness. Visualization of the changes in the feature maps of the ARN_3 layer of the model before and after the introduction of AttResNet is shown in Fig. 4.

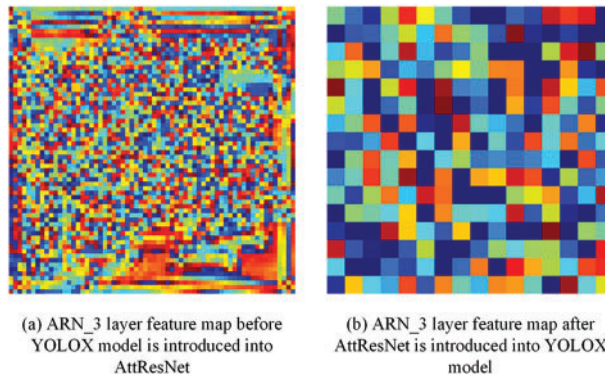


Figure 4: ARN_3 layer feature map visualization before and after the introduction of AttResNet in the YOLOX model backbone network

As can be seen from Fig. 4, the information acquired by the ARN_3 layer after the introduction of AttResNet into the backbone network of the YOLOX model is more condensed, the non-important information has been filtered, and the main information has been retained and aggregated, which effectively improves the model's ability to extract highly semantic information.

2.3 Target Detection Framework Backbone Partially Incorporates AttResNet Network

The proposed enhanced residual component AttResNet is fused in the backbone feature extraction part of the target detection network, where the residual network is commonly used because of its optimal performance. Theoretically, any target detection framework that uses ResNet can be replaced by AttResNet. As shown in Section 2.2, AttResNet has better feature extraction performance than traditional ResNet. To illustrate the optimal adaptability and performance of AttResNet in extracting tiny features, the backbones of AttResNet and the five main target detection frameworks (YOLOv3-SPP [23], YOLOX, Faster R-CNN [23], TDD-Net [24], Cascade R-CNN [25]) are fused to form a new target detection framework. The following is an example of how to use YOLOX-s to modify the backbone part. The network structure diagram is shown in Fig. 5.

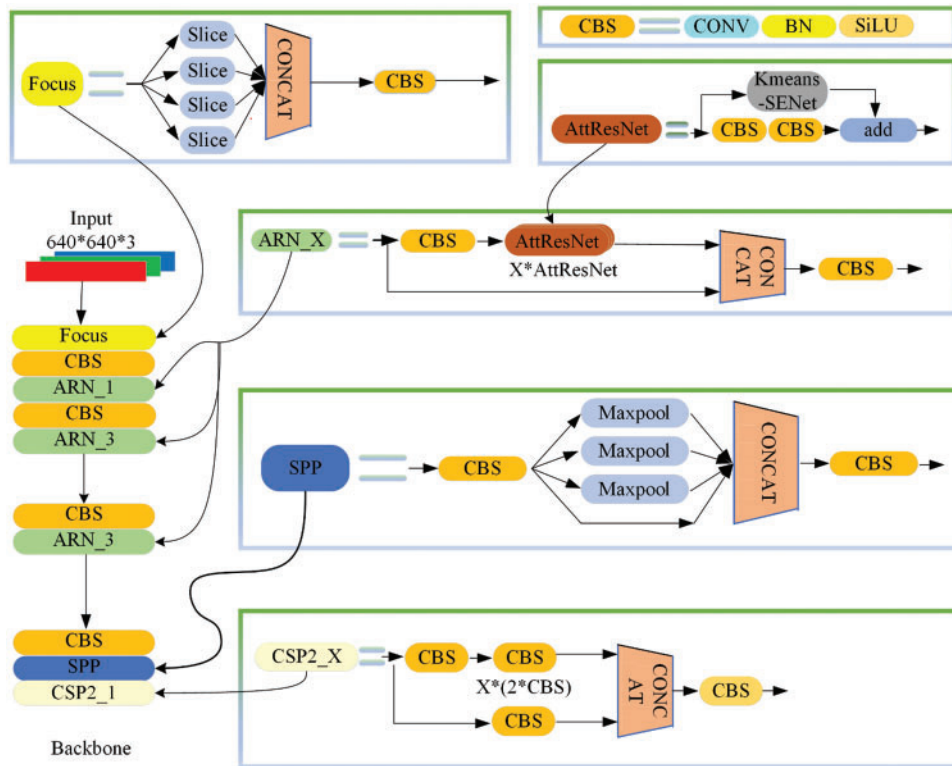


Figure 5: Combination structure of AttRest module and backbone

In Fig. 5, YOLOX mainly uses two data enhancement methods, Mosaic and Mixup, at the input side of the network. In the backbone part, the core change is that the Resunit component is replaced by the AttResNet component. The CSP structure will also be modified to a new structure with the AttResNet component. Other components, such as Focus, CBS, and SPP components remain unchanged. The Focus module was first proposed in YOLOv5 [26] and is usually used before the picture enters the backbone. CBS is the smallest component of the YOLOX network structure, which is composed of the triplet Conv+Bn+Silu activation function. SiLU is an improved version of Sigmoid and ReLU.

The AttResNet component borrows the residual structure from the ResNet network, allowing the network to be built deeply. This component is the most basic core unit for extracting high semantic features from the deep network. ARN_X component borrows the CSPNet network structure [27], and

it is composed of the CBS component and X AttResNet modules Concrete. The ARN_X component forms the main part of the backbone, and the backbone is deepened by stacking at different depths to obtain higher semantic features. The SPP component uses 1×1 , 5×5 , 9×9 , and 13×13 maximum pooling for multi-scale fusion to improve the scale invariance and reduce overfitting of images.

2.4 Target Detection Framework Backbone Partially Incorporates AttResNet Network

The AttResNet component is fused in the backbone of the target detection framework by the method in Section 2.3. The structure of YOLOX-AttResNet is shown in Fig. 6. It consists of 4 parts: input, backbone, neck, and detection head.

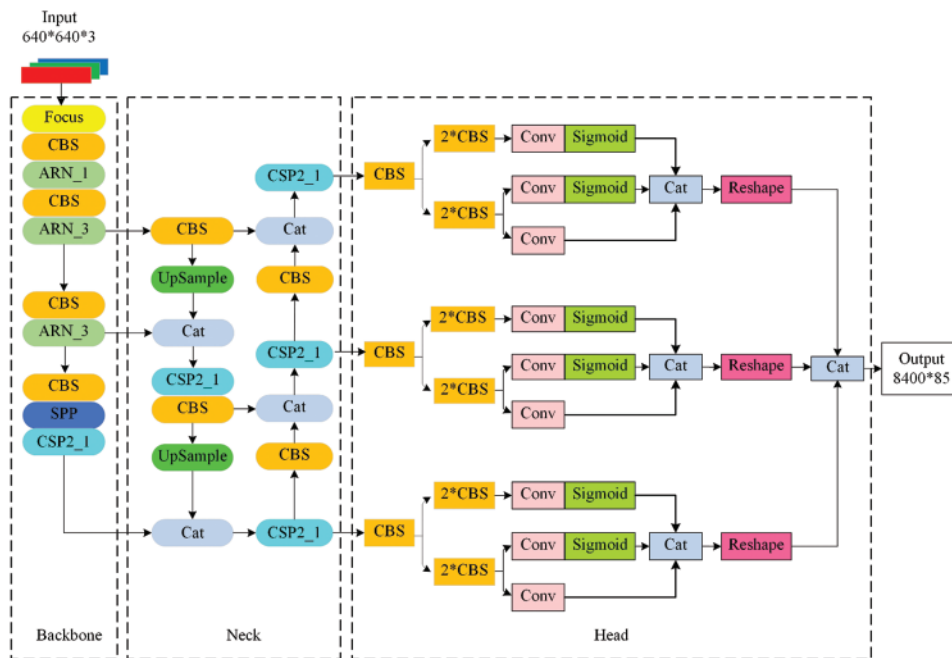


Figure 6: Complete network model diagram

The input part of the YOLOX-AttResNet framework uses Mosaic and Mxiup data enhancement methods to attenuate the drawbacks of insufficient data volume. The backbone feature extraction part uses the enhanced backbone described in Section 2.3, and the enhanced backbone is also applicable in other target detection frameworks. The neck uses the FPN structure for feature fusion to enhance feature extraction. FPN is top-down, and the high-level feature information is fused through up-sampling to obtain the feature map for the next prediction. In the output layer, YOLOX has many innovations, such as using a Decoupled Head structure with better performance, using Anchor Free mode instead of the previous Anchor mode, introducing SimOTA for fine-grained swiping in the label assignment, and using a Focal Loss network with more accurate computation and faster convergence.

3 Minor Defect Evaluation Tests in PCB Surface Inspections

In this section, further ablation experiments of the proposed YOLOX-AttResNet framework in typical tiny target PCB defects are carried out to validate the performance of AttResNet in detecting tiny targets of high-resolution low-contrast images.

3.1 Dataset Introduction and Experiment

In this work, we used the PCB surface defect dataset provided by Peking University for evaluation [24]. The PCB defect dataset published by the Human-Computer Interaction Open Laboratory of Peking University has 693 images with 3–5 defects on each image. The defects include six types: missing hole, mouse bite, open circuit, short, spur, and spurious copper. The defect images in the original dataset have high resolution. For such a small dataset, the data enhancement techniques are used before training. The images are then cropped into 600×600 sub-images, and a training and test set of 9920 and 2508 images are obtained, respectively. Most defect targets are tiny targets, and the defects and backgrounds are mixed to form a low contrast, which makes it challenging to accurately locate and classify defects.

3.2 Experimental Platform

The experimental setup involved in the experiments of this paper is shown in [Table 1](#).

Table 1: List of experimental setups

Device name	Causality
Deep learning modeling framework	PyTorch
Editorial platforms	Pycharm
Computer model	Intel® Core™ i5-12600KF 3.70 GHz
CPU memory	16 GB
Graphics card model	NVIDIA GeForce RTX 3070
GPU memory	8 GB
Computer system	Windows 10

All training and evaluation are conducted on the open-source toolbox PyTorch via the Pycharm platform on a computer with a 12th Gen Intel® Core™ i5-12600KF 3.70 GHz CPU and 16 GB of installed memory. An NVIDIA GeForce RTX 3070 GPU on Windows 10 (with 8 GB of memory) is used to run YOLOX-AttResNet and other target detection models involved in the ablation experiments. The experimental parameters of the model are set in [Table 2](#).

Table 2: Experimental setup

Parameter	Value
Iteration	300
Batch size	16
Initial learning rate	0.1
Min learning rate	0.0001
Optimizer	SGD
Momentum	0.937
Weight decay	5e-4
Learning rate decay type	COS
Thread	4

Because the model parameters affect the model performance, to avoid the influence of model parameters on the experimental results, the model parameters in Table 2 were kept constant for all the experiments involved in this paper. The parameter settings in Table 2 are the best parameter values summarized through several experiments, so they are used in all models to ensure objectivity.

3.3 Evaluation Indicators

Unlike traditional classification algorithms, the relevant metrics to evaluate the effectiveness of the convolutional neural network target detection model are mainly mAP and FPS. AP denotes the area under the Precision-Recall curve, and mAP is the average of the APs of multiple categories, with the range of $[0, 1]$; AP evaluates the local performance of the learning model of each category, while mAP evaluates the overall performance of the learning model of all categories. Since the target detection task in this experiment is multi-category, mAP is selected as the evaluation metric; FPS, i.e., the number of images that can be processed in one second, and the comparison are conducted on the same hardware.

3.4 PCB Surface Micro Defect Definition and Ablation Experiment

According to the definition of SPIE, a small target is an area less than 80 pixels in a 256×256 image, i.e., less than 0.12% of 256×256 is a small target. As shown in Fig. 7, there are small, medium, and large defect detail maps in the high-pixel image (3056×2464).

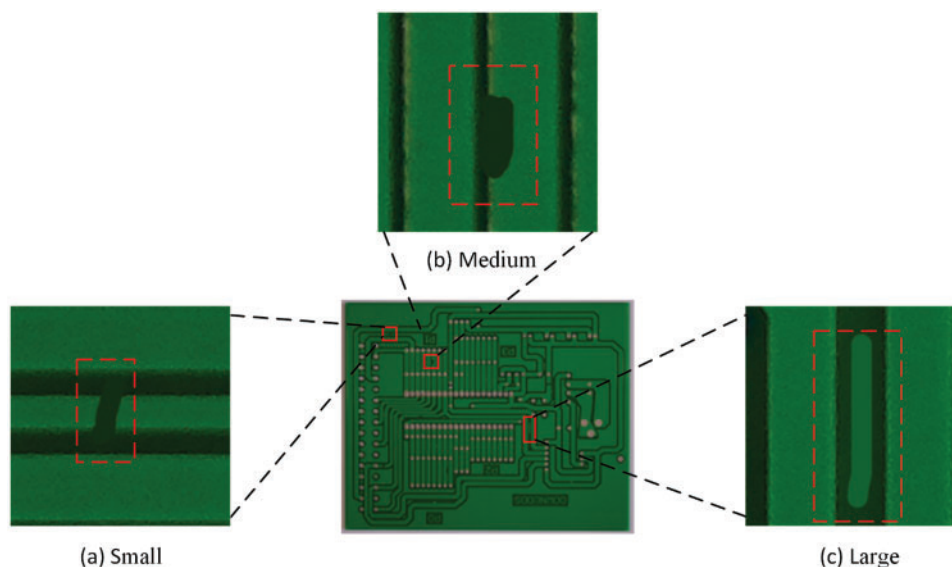


Figure 7: Shows the defect details section. (a) small size defects, (b) medium size defects, and (c) large size defects

The improvement point of the model in this paper mainly lies in the introduction of the AttResNet module in the backbone network, which is novel in that it modifies the traditional ResNet network to map branches directly, and replaces it with a channel attention module using K-means optimization as a means of attenuating the direct propagation of redundant features to achieve the enhancement of the detection capability of tiny targets. To verify the performance of the AttResNet-YOLOX model proposed in this paper in detecting tiny defects on PCB surfaces, a full ablation study is done on the PCB surface defects dataset provided by Peking University, and the experimental results are recorded in Table 3.

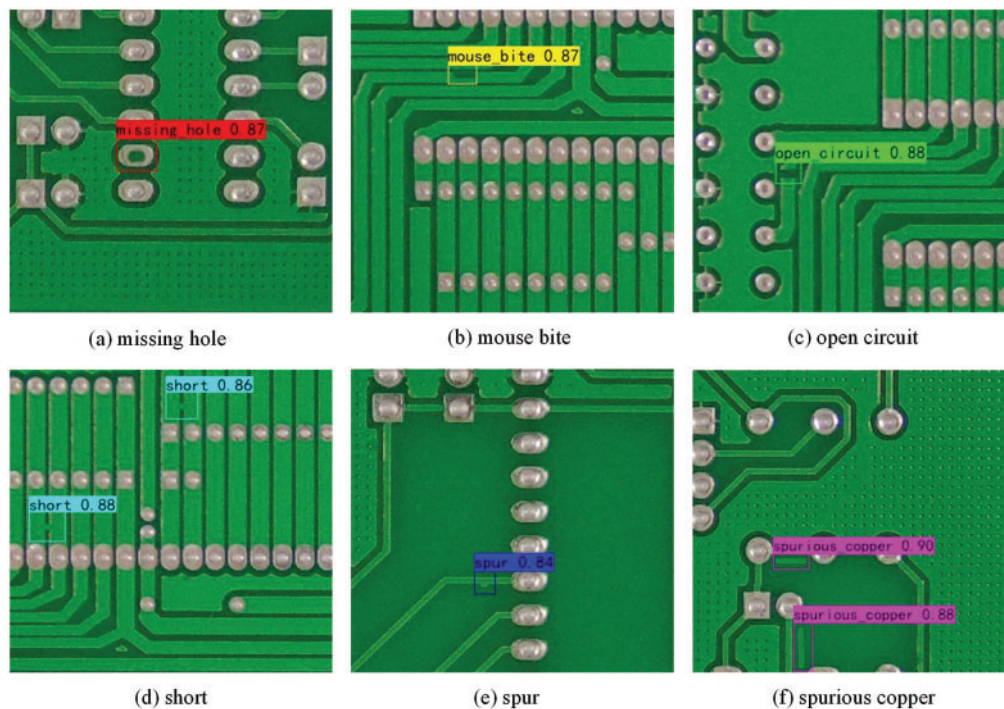
Table 3: Ablation results of PCB surface defect data set

Algorithms	mAP (%)	mAPS (%)	mAPm (%)	mAPI (%)
YOLOX	96.24	75.29	84.67	98.35
+SENet	96.37	76.32	85.37	98.87
+AttResNet	98.45	82.96	87.58	99.56

Analysis of [Table 3](#) shows that the original YOLOX model mAP is 96.24% when adding SENet in the direct mapping branch of the base ResNet mAP is 96.37, the change is small, indicating that the simple addition of the Attention Mechanism module cannot bring about a significant improvement in detection accuracy when adding the AttResNet module at the same location of the base ResNet mAP is 98.45, the detection accuracy is improved by 2.21%, which is a large improvement, indicating that the AttResNet module effectively improves the feature extraction performance of the backbone network. In addition, in the mAPs indicator, when the AttResNet module is added to the backbone network, the mAPs are improved by 7.67%, indicating that the contribution of the AttResNet module is more obvious on small targets.

3.5 Performance Evaluation of YOLOX-AttResNet

A high pixel image (3056×2464) with all defects (total 16 defects) and 6 individual defective normal pixels (600×600) are input into YOLOX and YOLOX-AttResNet networks for testing. The detection results of 600×600 normal pixel images are shown in [Fig. 8](#), and that of 3056×2464 high pixels are shown in [Fig. 9](#).

**Figure 8:** Six kinds of defect detection results of small size (200×200) images

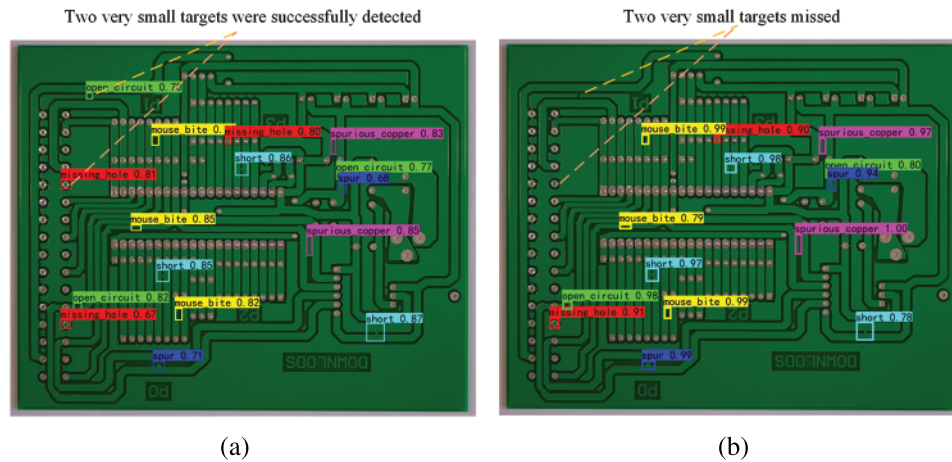


Figure 9: YOLOX-AttResNet and YOLOX detection results in 3056×2464 pixel images. (a) YOLOX-AttResNet test result (b)YOLOX test result

As shown in Fig. 8, in the low-pixel single targets, YOLOX and YOLOX-AttResNet networks successfully detected all defects in all images, indicating that they have optimal detection performance for low-pixel single targets. As shown in Fig. 9, all 16 targets are detected by YOLOX-AttResNet, and two small target defects are missing by YOLOX. It indicates that YOLOX-AttResNet has better performance in the small targets with high-pixel multi-target low-contrast.

Using the Class Activation Mapping (CAM) method, the detection results of YOLOX and YOLOX-AttResNet models are visualized in features. This makes it easy to observe which areas of the image in the network focus and reinforce before the target is identified and located after the backbone. In addition, it is easy to analyze and compare different models on the same target, and the feature heat map is shown in Fig. 10.

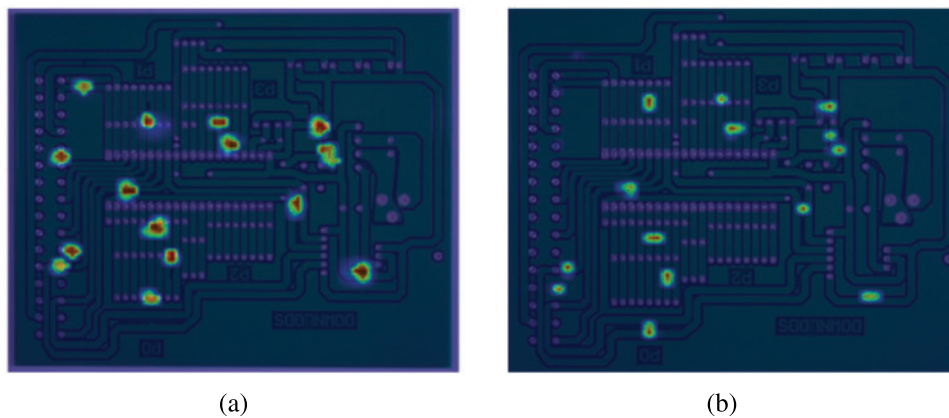


Figure 10: Comparison of YOLOX-AttResNet and YOLOX heat map in the backbone section. (a) Network concern region and degree in YOLOX-AttResNet detection; (b) Network concern region and degree in YOLOX detection

From Fig. 10, it can be seen that the YOLOX-AttResNet network modifies the basic ResNet module in the backbone into an AttResNet module, which focuses on all targets to enhance the

features. The AttResNet module can enhance the performance of extracting small features in the network.

To measure whether a model is robust and reliable, it is necessary to calculate the size of the area covered by the precision-recall curve. Fig. 11 shows the precision-recall (PR) curves for six defects on a PCB defect image at IoU = 0.5.

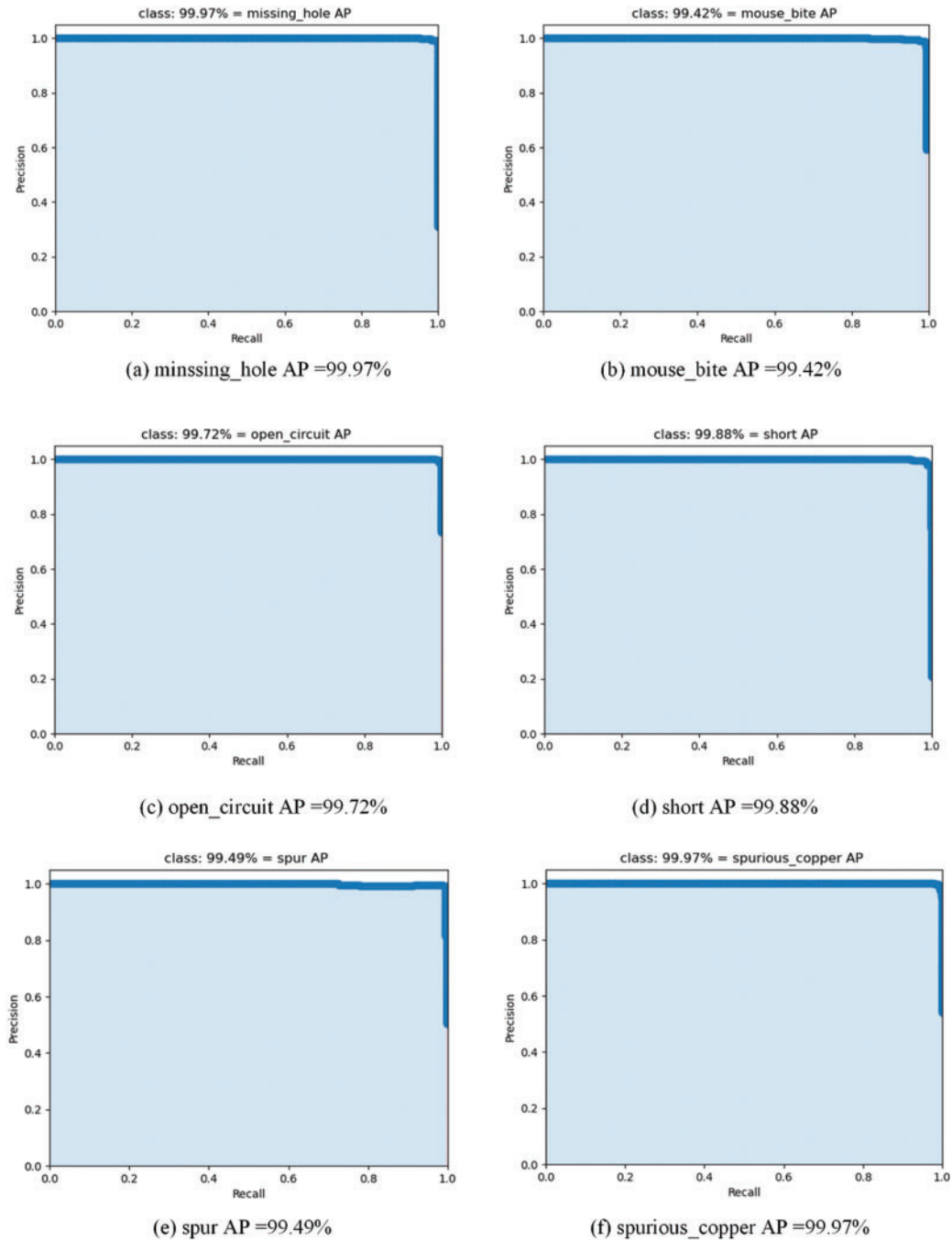


Figure 11: The mAP@0.5PCB defect accurate recall (PR) curve

The exact recall curves for each of the six defective targets are calculated in Fig. 11. It can be seen that the area under the PR curves of the six targets is nearly 1, indicating that the YOLOX-AttResNet network has high detection accuracy and recall rate for all targets, with optimal robustness.

3.6 Detection Performance Comparison of YOLOX-AttResNet and Other 5 Excellent Algorithms

To verify the excellence and good combination of the proposed AttResNet module in feature extraction, it is combined with the backbone part of the current best 5 target detection framework. Comparative experiments are carried out on the PCB surface defect dataset for accuracy and detection speed metrics. The test results are shown in Table 4.

Table 4: Comparison of detection results of different target detection frameworks fused with the AttResNet module

Algorithms	mAP@0.50 (%)	Speed (bs = 1, FPS)	Dim1 UP (mAP)	Dim2 UP (mAP)
YOLOv3	95.50	22	–	–
YOLOv3 + AttResNet	96.90	19		1.4%↑
Faster R-CNN	94.44	15	1.06%↓	–
Faster R-CNN + AttResNet	94.68	13		0.24%↑
TDD-Net	95.90	62	0.40% ↑	–
TDD-Net + AttResNet	96.10	53		0.2%↑
Cascade R-CNN	96.20	9	0.70% ↑	–
Cascade R-CNN + AttResNet	96.30	8		0.1%↑
YOLOX	97.24	42		–
YOLOX + AttResNet	98.45	36	2.95%↑	1.21%↑

Table 4 shows that when the PCB defect dataset published by the Open Laboratory for Human-Computer Interaction of Peking University is used as the model training and testing data source, the performance of five excellent algorithms (including YOLOv3, YOLOX, Faster R-CNN, TDD-Net, and Cascade R-CNN) is compared in two dimensions. Among them, TDD-Net is a network model designed for PCB surface defect data sets. In the first dimension, using YOLOv3 as the benchmark model, the YOLOX+AttResNet algorithm has the largest mAP improvement of 2.95% when comparing the mAP values obtained from the other five models tested. In the second dimension, the five excellent algorithms with the AttResNet module added are compared with the original algorithm separately, and all have improved accuracy, with a maximum of 1.40% mAP detection accuracy improvement. YOLOX-AttResNet shows the best performance to reach the final 98.45% mAP. Although the detection speed is reduced by 4.5%, it fully meets the requirements of high detection accuracy and real-time performance for small targets of PCB surface defects. According to introducing an attention mechanism, the information that is more critical to the current task is focused to reduce the attention to other information, or even filter out irrelevant information. The problem of information overload can be solved and the accuracy of task processing can be improved.

4 Conclusion and Statements

With the development of modern electronic products to micro-small, high-density, thin, and lightweight, PCB board circuitry to carry electronic components has become more dense and complex, which presents new challenges for PCB board circuit detection. In this paper, six common circuit defects such as missing holes, mouse bite, open circuits, short, spur, spurious copper, etc. are detected on the surface of high-pixel large-size PCB boards, and the YOLOX-AttResNet model is designed for the detection of circuit defects on the surface of PCBs, and experiments are conducted on the surface circuit defects dataset of PCB. The experimental results on the surface circuit defect dataset show that the AttResNet module proposed in this paper is reliable and efficient, and when combined with the five best existing target detection frameworks (YOLOv3, YOLOX, Faster R-CNN, TDD-Net, and Cascade R-CNN), the detection accuracies are improved compared to the original model, where. The best performance is achieved when fused with the YOLOX model, reaching a maximum accuracy of 98.45% and a detection speed of 36 FPS (Frames Per Second), which indicates that the AttResNet module has a good combination and excellent feature extraction performance for small features. The YOLOX-AttResNet model proposed in this paper meets the accuracy and real-time requirements for detecting small defects in PCB surface circuits, especially small defects that are easy to miss, such as missing holes and short defects.

AttResNet module has significant advantages in extracting features of tiny defects on the PCB surface, which can increase the amount of information on weak features, eliminate the redundant information similar to the background, and retain the key features of small targets; however, at the same time, the module also comes with some negative impacts, such as increasing the complexity of the model, and the calculation of the attenuation of the redundant information in each base AttResNet module, which aggravates the burden of the model, and the efficiency of the operation is reduced.

Acknowledgement: Thank you to our researchers for their collaborations and technical assistance.

Funding Statement: This work was supported by the National Natural Science Foundation of China (No. 61976083); Hubei Province Key R&D Program of China (No. 2022BBA0016).

Author Contributions: Xinyu Hu and Defeng Kong completed the work and contributed to the writing of the manuscript. Xiyang Liu and Junwei Zhang conceived the project and provided the research methodology. Daode Zhang conducted the survey and data management for the project. All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Availability of Data and Materials: The data that support the findings of this study are openly available in (AttResNet) at <https://github.com/jackong180/AttResNet>.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] X. Hu, D. Kong, X. Liu, J. Zhang and D. Zhang, "FM-STDNet: High-speed detector for fast-moving small targets based on deep first-order network architecture," *Electronics*, vol. 12, no. 8, pp. 1–15, 2023.
- [2] X. Liao, J. Zhang and S. Lu, "PCB defect detection methods incorporating shallow features and attention mechanisms," *Computer Integrated Manufacturing System*, vol. 28, pp. 1–17, 2023.

- [3] M. Abd Al Rahman and A. Mousavi, "A review and analysis of automatic optical inspection and quality monitoring methods in electronics industry," *IEEE Access*, vol. 8, pp. 183192–183271, 2020.
- [4] R. Chen, Z. Zhan, X. Hu, X. Xu, D. Cai *et al.*, "Multi-attention faster RCNN-based defect detection of printed circuit boards under noise interference," *Journal of Instrumentation*, vol. 42, no. 12, pp. 167–174, 2021.
- [5] Z. Liu and B. Qu, "Machine vision based online detection of PCB defect," *Microprocessors and Microsystems*, vol. 82, pp. 103807, 2021.
- [6] Y. Wu, L. Zhao, Y. Yuan and J. Yang, "Research status and prospects of PCB defect detection algorithms based on machine vision," *Journal of Instrumentation*, vol. 43, no. 8, pp. 1–17, 2022.
- [7] I. C. Chen, R. C. Hwang and H. C. Huang, "PCB defect detection based on deep learning algorithm," *Processes*, vol. 11, no. 3, pp. 775, 2023.
- [8] H. Zhang, L. Jiang and C. Li, "CS-ResNet: Cost-sensitive residual convolutional neural network for PCB cosmetic defect detection," *Expert Systems with Applications*, vol. 185, pp. 115673, 2021.
- [9] J. H. Park, Y. S. Kim, H. Seo and Y. J. Cho, "Analysis of training deep learning models for PCB defect detection," *Sensors*, vol. 23, no. 5, pp. 2766, 2023.
- [10] J. Wu, Y. Cheng and J. Shao, "Study on PCB defect detection based on improved YOLOv4 algorithm," *Journal of Instrumentation*, vol. 42, no. 10, pp. 171–178, 2021.
- [11] Y. Yang and Y. Xia, "Precision improvement of deep learning algorithm for PCB defect detection," *Journal of Electronic Measurement and Instrumentation*, vol. 37, pp. 1–10, 2023.
- [12] J. Hu, S. Jia and S. Ma, "PCB surface defect detection algorithm based on improved cascade faster RCNN," *Instrumentation Technology and Sensors*, vol. 51, no. 7, pp. 106–110+126, 2022.
- [13] X. Wang, J. S. Gao, B. J. Hou, Z. S. Wang, H. W. Ding *et al.*, "A lightweight modified YOLOX network using coordinate attention mechanism for PCB surface defect detection," *Ieee Sensors Journal*, vol. 22, no. 21, pp. 20910–20920, 2022.
- [14] S. Y. Xia, F. Wang, F. Xie, L. Huang, Q. Wang *et al.*, "An efficient and robust target detection algorithm for identifying minor defects of printed circuit board based on PHFE and FL-RFCN," *Frontiers in Physics*, vol. 9, pp. 1–9, 2021.
- [15] J. Tang, S. Liu, D. Zhao, L. Tang, W. Zou *et al.*, "PCB-YOLO: An improved detection algorithm of PCB surface defects based on YOLOv5," *Sustainability*, vol. 15, no. 7, pp. 5963, 2023.
- [16] B. Hu and J. Wang, "Detection of PCB surface defects with improved faster-RCNN and feature pyramid network," *IEEE Access*, vol. 8, pp. 108335–108345, 2020.
- [17] B. Chen and Z. Dang, "Fast PCB defect detection method based on fasterNet backbone network and CBAM attention mechanism integrated with feature fusion module in improved YOLOv7," *IEEE Access*, vol. 11, pp. 95092–95103, 2023.
- [18] Z. Ling, A. Zhang, D. Ma, Y. Shi and H. Wen, "Deep siamese semantic segmentation network for PCB welding defect detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, 2022.
- [19] K. He, X. Zhang and S. Ren, "Identity mappings in deep residual networks," in *Computer Vision–ECCV 2016*, Amsterdam, The Netherlands: Springer, pp. 630–645, 2016.
- [20] M. Ahmed, R. Seraj and S. M. S. Islam, "The k-means algorithm: A comprehensive survey and performance evaluation," *Electronics*, vol. 9, no. 8, pp. 1295, 2020.
- [21] J. Hu, L. Shen and G. Sun, "Squeeze-and-excitation networks," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 7132–7141, 2018.
- [22] K. He, X. Zhang, S. Ren, J. Sun, "Identity mappings in deep residual networks," in *14th European Conf. on Computer Vision (ECCV)*, Amsterdam, Netherlands, 2016.
- [23] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.
- [24] R. Ding, L. Dai, G. Li and H. Liu, "TDD-net: A tiny defect detection network for printed circuit boards," *CAAI Transactions on Intelligence Technology*, vol. 4, no. 2, pp. 110–116, 2019.
- [25] Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving into high quality object detection," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Salt Lake City, Utah, USA, pp. 6154–6162, 2018.

- [26] X. Zhu, S. Lyu and X. Wang, "TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios," in *Proc. of the IEEE/CVF Int. Conf. on Computer Vision*, Montreal, BC, Canada, pp. 2778–2788, 2021.
- [27] C. Y. Wang, H. Y. M. Liao, Y. H. Wu, P. Y. Chen, J. W. Hsieh *et al.*, "CSPNet: A new backbone that can enhance learning capability of CNN," in *Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops*, Montreal, BC, Canada, pp. 390–391, 2020.