IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 31 March 2024, accepted 23 April 2024, date of publication 30 April 2024, date of current version 9 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3395499

RESEARCH ARTICLE

Defect Detection Model of Printed Circuit Board Components Based on the Fusion of Multi-Scale Features and Efficient Channel Attention Mechanism

WENBIN CHEN[1], HONGCHAO ZHAO[1], AND ZHENG WANG[2]
[1]School of Electrical Engineering, Chongqing University of Science and Technology, Chongqing 401331, China
[2]Faculty of Engineering, Mie University, Tsu, Mie 514-8507, Japan

Corresponding author: Wenbin Chen (chenwb@cqust.edu.cn)

This work was supported in part by the Science and Technology Research Program Project of Chongqing Municipal Education Commission under Grant KJZD-K202203601, in part by the School-Level Project of Chongqing University of Science and Technology under Grant ckrc2020009, and in part by the Project of Deep Learning-Based Online Detection of Assembly Defects in Printed Circuit Board Components under Grant 19230101029.

ABSTRACT The detection of defects in printed circuit board (PCB) components is crucial to the quality of PCB. Issues such as blurred details, complex and varied backgrounds, and inadequate recognition of PCB components lead to poor detection accuracy. To address these challenges, this paper introduces a PCB component defect detection model (MSF-ECANet) based on multi-scale features and efficient channel attention networks. Firstly, to address the challenge of unclear information regarding intricate features in deep networks, Residual Nets (ResNet) and Multi-Scale Feature Pyramid Networks (FPN) are integrated. This fusion tackles the issue of vanishing gradients, expands the model's receptive field, and optimizes the model's proficiency for recognizing PCB components. Secondly, to improve the recognition rate of PCB component detection, Efficient channel attention networks (ECA-Net) are used to assign different weights to the PCB background and foreground channels to segment the background and foreground. Lastly, a dichotomous K-means algorithm is used to obtain the optimal anchor size that is closer to the ground truth size, so as to improve the sensitivity of the model to small target detection. When compared to SSD, YOLOv3, YOLOv5, YOLOx and Faster R-CNN, the experimental results show that the model proposed in this paper improves 1.41%, 7%, 4.17%, 5.47% and 8.33% in accuracy, respectively. Furthermore, the improved network exhibits superior convergence compared to the original network. Therefore, the MSF-ECANet model presented in this paper is more suitable for industrial applications of PCB component defect detection.

INDEX TERMS PCB component defects, ResNet, FPN, ECA-Net, dichotomous k-means.

I. INTRODUCTION

The printed circuit board serves as the structural foundation for electronic components., The presence of missing or incorrect components may cause circuit short circuits within the PCB, potentially resulting in component ignition or blast, eventually, leading to economic losses for the enterprise [1]. Consequently, numerous scholars have been devoted in PCB defect detection. While most researchers are targeted at PCB bare board defects, such as internal line detection, scratch defects, solder joints, only a minority of scholars have concentrated on PCB components, including: resistors, capacitors, power supply slots. So far, there are five types of inspection methods for PCB component defects: traditional manual inspection, Automatic Optic Inspection (AOI), image segmentation inspection [2],

The associate editor coordinating the review of this manuscript and approving it for publication was Tallha Akram.

machine learning inspection [3], and deep learning inspection [4].

Manual inspection methods involve direct visual inspection or checking components using a magnifying glass [5]. The advantage is that they are low cost and do not require additional physical resources such as fixtures, but they are inefficient and low accuracy.

With the development of PCB integration, AOI technology [6], [7], [8] has been integrated. AOI offers numerous advantages such as integrated design, programmable control, and a high level of durability. But AOI is not perfectly suited for the wide array of PCB components and cannot conduct rapid, efficient, real-time in-line inspection on PCB assembly lines.

Image segmentation techniques have been proposed to solve the problem of real-time PCB defect detection and the detection of a single type of defect [9], [10], [11]. Although the image segmentation technique can meet the accuracy and real-time of PCB defect detection in certain scenarios, it requires a large number of experiments to debug parameters and template matching, as well as high quality PCB datasets to produce the desired results.

With a series breakthroughs of hardware technology, including high-performance computers, graphics processors, tensor processors, and high-bandwidth memory, deep learning has been developed significantly, Researchers have deepened neural network layers to extract clearer information regarding the semantic and detailed features of PCB images. Currently, deep learning-based PCB defect detection methods are mainly one-stage algorithms (such as You Only Look Once, YOLO [12], Single Shot MultiBox Detector, SSD [13]), and two-stage algorithms (like Faster R-CNN [14]). For instance, some researchers have employed one-stage detection algorithms for PCB defect detection, Deng et al. [15] combined CNN technology with AOI technology to locate PCB line defects using difference operations to compensate for the detailed information in PCB images at the feature extraction stage, but the method can produce pseudo defects due to contaminants like dirt and dander on the PCB surface. Zhang et al. [16] introduced a sliding window method in CNN model to locate PCB defects, thereby improving the detection accuracy and reducing the complexity of traditional vision methods, but resulted in a long detection time due to the complexity of the model's architecture. Lin et al. [17] pioneered to use the YOLO model to detect PCB assembly capacitors, which had certain advantages in terms of detection speed, but because the PCB dataset is small, the model's generalization capability is poor. He and Yu [18] added a long-range global attention mechanism based on a fine-grained spatial domain to YOLOv4 to enrich of feature information, but the accuracy only reached 91%. Some people also use Two-Stage detection algorithm to detect PCB defects. For example, Hu and Wang [4] realized the feature fusion of different sizes of target information through up-sampling method and jump layer connection to solve the problem of target detection range of PCB tiny defects and improve the model accuracy. Sun et al. [19] integrated DCN method into ResNext101 network and embedded NAS-FPN network in RPN network to improve the accuracy and speed for PCB high-density defect detection. Zhang et al. [20] proposed the IDD-Net network for defect detection on three industrial datasets, and the accuracy of the proposed IDD-Net network achieved 95.9% mAP for PCB bare board defect detection, 99.5% mAP for aluminum surface defect detection, and 79.6% mAP for steel surface defect detection. Although this method can be adapted to a variety of data sets, the average accuracy of defect detection still needs to be improved. A comprehensive summary and analysis of these deep learning methods are presented in Table 1, offering insights into their performance in PCB defect detection.

TABLE 1. Characterization table for deep learning in PCB defect detection (○: Excellent, Δ: Fair, ×: Poor).

Characteristics	One-stage		Two-stage
	YOLO	SSD	Faster R-CNN
Accuracy	Δ	Δ	○
Speed	○	○	Δ
Applicability	○	○	○
Volume of Calculations	×	○	×
Cost	Δ	Δ	Δ
Feature Extraction Capability	Δ	×	○
Modelling Difficulty	Δ	Δ	Δ

In view of the excellent performance of two-stage network detection, this paper proposes MSF-ECANet model to detect PCB component defects. The main contributions of this study are outlined below:

1. This paper proposes a method that fuses ResNet and FPN to transform the original Faster R-CNN model from single-scale feature information extraction to multi-scale feature information extraction. The method addresses the issue of vanishing gradient and enlarges the receptive field, which effectively improves the model's ability to capture the features of PCB components, and then improves the classification accuracy of the model.

2. The high similarity between the foreground (PCB components) and the background poses a challenge to the effective feature segmentation of the model. To solve this problem, this paper proposes a feature segmentation method based on ECA-Net, which effectively segments the background and foreground by assigning different weights to the PCB background and foreground channels. Thus the model's recognition rate of PCB components is improved.

3. In PCB images defects detection, dense small targets are prone to be missed. To solve this problem, this paper proposes an anchor frame optimization method based on dichotomous K-means algorithm. The method obtains the optimal anchor box size by using the dichotomous K-means algorithm in the Faster R-CNN model, which improves the sensitivity of the model in detecting dense small targets.

The paper is structured as follows: Section II briefly reviews the related work on faster R-CNN networks; Section III elaborates on the MSF-ECANet network framework, multi-scale feature fusion, ECA-Net network and anchor size design optimization; Section IV presents experimental validation and comparative analysis of the MSF-ECANet model; Section V concludes the research of this paper.

II. RELATED WORK

In recent years, the rapid development of deep learning has led to a shift in target detection from the traditional method of manually designing features to the deep learning method. Compared with traditional methods, deep learning target detection has the advantages of strong stability, high accuracy and high efficiency. Currently, deep learning target detection can be divided into two main approaches: One-stage detection algorithm and Two-stage detection algorithm. The One-stage algorithm directly extracts features from the input images to predict both classification and localization. These algorithms, such as YOLO series and the SSD algorithm, are known for their simplicity, speed, and effectiveness in scenarios where targets are prominent. The Two-stage algorithm extracts feature through the candidate region box, and divides the target detection into two stages: firstly, it extracts the features from the input image and generates candidate box, which are then refined through subsequent adjustments to improve their positioning accuracy. Lastly, a classifier is used for image classification and localization. Typical Two-stage algorithms are the R-CNN (Regions with CNN Features) family networks.

R-CNN is regarded as a typical representative of Two-stage target detection algorithm [21], but it suffers from three problems: 1) the R-CNN network can't be trained individually, it must be divided into multiple steps for training, which results in a long training; 2) the cropped candidate frames need to be compressed or pulled up, this step leads to image distortion; 3) the network requires a large amount of memory to store the candidate frames. To address these issues, the Fast R-CNN [22] algorithm was introduced, enabling end-to-end training, improving training speed, and eliminating image distortion. However, its accuracy still needed improvement. Therefore, Ren et al. [14] proposed Faster R-CNN model, achieving complete end-to-end target detection with better performance in terms of accuracy and speed.

As shown in Figure 1, the Faster R-CNN consists of three parts: feature extraction network (FEN), region proposal network (RPN), and classification network. The Faster R-CNN accepts image of arbitrary size, initially resizing them to 600×600 before forwarding them through the detection network. Within this network, the feature extraction network generates a feature map. Subsequently, the RPN module generates candidate suggestion frames and maps the generated candidate suggestion frames to the feature map to obtain the feature map containing the suggestion frames. Lastly, the feature map is passed through the RoI (Region of Interest) pooling layer and the fully connected layer to complete the classification prediction and regression prediction. The regression prediction part is used to adjust the position and the size of the candidate proposal boxes generated in the RPN, and the classification prediction part is used to obtain the model prediction results.

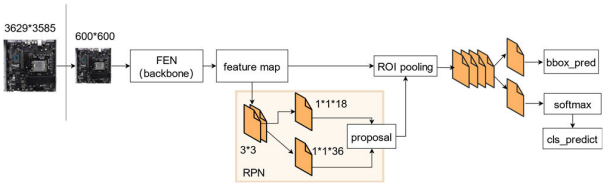

FIGURE 1. The original Faster R-CNN detection framework.

Based on the superior accuracy of Faster R-CNN, a lot of researchers have applied the Faster R-CNN model to PCB detection. In [23], K-means clustering was used to design reasonable anchor box size, and online hard example mining technique was used to solve the problem of sample imbalance and to optimize anchor box size. However, the boxes generated by K-means algorithm are prone to missed detection. Hu and Wang [4] integrated ResNet50 [24] as a feature extraction network in the Faster R-CNN network, and introduced the GARPN network and residual units of ShuffleNetV2 to reduce the training time by reducing the number of anchor points. Nevertheless, the method suffers from PCB feature loss during the mapping process. Zhu et al. [25] fused multiscale features and deformable convolutional networks to improve the model's ability to extract PCB image feature information. However, due to the limited diversity of PCB dataset, the model exhibits poor generalization ability. Jiang et al. [26] proposed a RAR-SSD model, which combines a multi-scale network with SSD network to solve the problem of defect leakage and false detection of PCB bare board. The RAR-SSD model enhances the receptive field and attention mechanism module, and uses the different weights of different features in different channels in PCB images to provide a wider range of effective focusing features. It uses a lightweight model with few parameters. However, the RAR-SSD model simultaneously employs both pixel-level and channel-level attention mechanisms, leading to a significant increase in computation. Additionally, there is a certain degree of PCB image data loss when fusing features extracted from the RAR-S.

The studies presented by Hu and Wang [4], Ren et al. [14], and Zhang et al. [20] show that the detection accuracy of the original Faster R-CNN model is higher than that of the YOLO [12] model and SSD [13] model in the small object detection task. So, in our work, the Faster R-CNN network framework is used as the basic model of the improved algorithm to detect the small target in PCB component.

III. PCB COMPONENT DEFECT DETECTION ALGORITHM

A. MSF-ECANET DETECTION NETWORK

The Faster R-CNN network framework uses VGG16 as the backbone network for feature extraction, and the weight

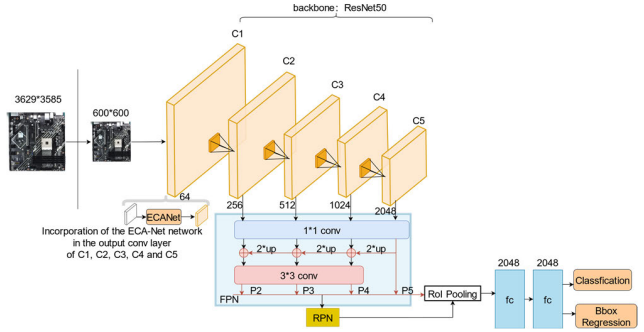

FIGURE 2. MSF-ECANet network structure.

update of each convolutional layer needs to be carried out by backpropagation during the training process. However, as the number of network layer increases, the problem of vanishing gradient becomes more significant, leading to a decrease in detection accuracy. To solve this problem, the ResNet network [24] introduced a residual module. In addition, ResNet preserves better feature information of the original image. Therefore, in this paper, instead of VGG16 network, ResNet network is used. However, the ResNet50 network uses a single-scale feature layer to construct the feature map, potentially leading to the missed detection when extracting feature information for PCB components. Therefore, in this paper, ResNet50 network and FPN network are fused. Additionally, to solve the problem that PCB components are highly similar to the background information features, the ECA-Net network is integrated into the output conv layers of C1, C2, C3, C4, and C5 of ResNet50 network. Our proposed MSF-ECANet network structure is shown in Figure 2.

B. MULTI-SCALE FEATURE FUSION

Since only ResNet50 is used as the backbone feature extraction network of Faster R-CNN, and the features extracted by its C5 layer are used to construct the feature map, it has poor ability to detect PCB components and extract feature information. Therefore, in this paper, FPN network is utilized to fuse the high-level semantic information with the shallow detail information to improve the model's target recognition and feature information extraction ability for different sizes of tiny dense PCB components. FPN network comprises three parts: feature extraction network, feature information fusion, and multi-scale merging.

The feature extraction network has C1, C2, C3, C4, and C5 layers, with each layer containing a different number of conv layers inside to extract the image feature information. Between the adjacent layers, the size of the image is down sampled by an integer multiple of two. The output image of each layer provides a source of feature information for the FPN network. Feature information fusion takes the image features extracted from the deeper layers in ResNet50, performs a 2-fold up sampling operation by the nearest neighbor interpolation algorithm to ensure the uniformity of the height and width of the newly fused image feature information, then,

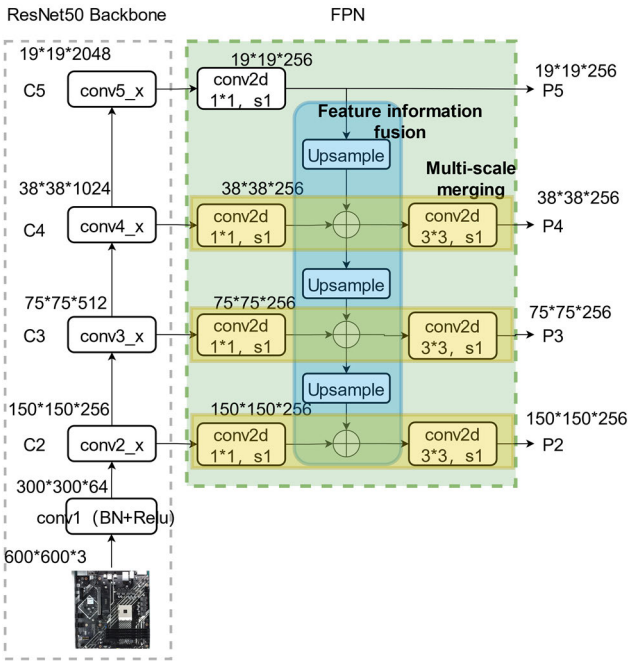

FIGURE 3. FPN network structure.

a new feature map layer is formed. This layer will then be merged with the shallow feature information in a multiscale merger. Multi-scale merging fuses the results of the feature information fusion part with the PCB image feature information output from each layer of the feature extraction network. Then, 3×3 conv layers are used to perform convolution operation on each fusion result to eliminate the aliasing effect generated by the up-sampling method during the feature information fusion process. The final feature map is generated and labeled as P2, P3, P4, and P5. The feature map of the RoI mapped to the corresponding layers of FPN [27] is calculated by (1).

$$k = k_0 + \log_2 \sqrt{wh}/600 \tag{1}$$

where $k_0 = 4$, w is the RoI width, h is the RoI height, 600 is the image size.

C. ECA-NET

For the characteristics of PCB components, background and foreground information may highly overlap, such as jumper04p (Figure 4(a)) and Jumper10p (Figure 4(b)). During the process of feature information extracting, the limitation of the convolutional layer structure can result in a weak expression ability for the feature information needed for target detection of different parts. This can ultimately lead to missed detections of PCB components, thereby reducing the overall accuracy.

Figure 5 illustrates the process of the fused ECA-Net network with ResNet50 network. The ECANet [28] extracts the feature information of the detected target by fusing the channel information within the local receptive field to

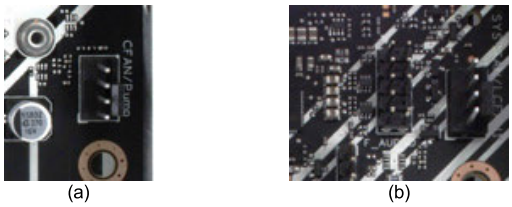

FIGURE 4. PCB components (components and background are highly similar). (a) Jumper04p; (b) Jumper10p.

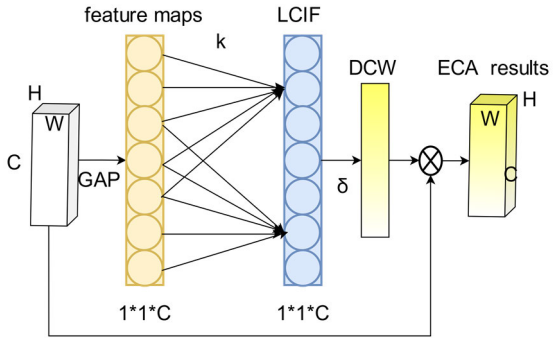

FIGURE 5. ECA-Net network structure.

improve the characterization capability of the network. The network structure uses parallel computation to avoid feature dimension reduction and improve the detection efficiency of missing PCB components. In addition, by employing shared weights, the ECANet minimizes model parameters and accelerates model convergence. The network structure reduces the complexity of the model by enhancing the interaction of information between channels in the PCB image. The integration of this network enhances the dependency between the channels of the PCB feature map, so that the PCB defect detection model pays more attention to the foreground channel feature information and suppresses the PCB background feature information, which improves the ability to express the useful feature information of the PCB foreground.

In ECANet, the input PCB feature maps (H*W*C) are subjected to Global Average Pooling (GAP) without dimensionality reduction to produce a feature maps of size 1∗1∗C. Next, the adaptive k-value is computed by (2). Subsequently, each channel and its k adjacent channels are obtained by one-dimensional convolution to obtain local cross-channel interaction features (LCIF). Lastly, the Sigmoid (δ) function is utilized to obtain the different channel weights (DCW). By fusing the channels with different weights with the channels of the input image, the weight information of each channel is finally obtained (ECA results in Figure 5), and different colors in the figure indicate different weights.

$$k = \psi(c) = \left| \frac{\log_2(c)}{\gamma} + \frac{b}{\gamma} \right|_{\text{odd}} \tag{2}$$

where $|t|_{\text{odd}}$ denotes the nearest odd number of t, $t = 2$, $b = 1$, and b is the number of channels.

D. OPTIMIZE ANCHOR SIZE

In the PCB dataset used in this paper, due to the presence of multiple components of different sizes, the target detection using Faster R-CNN requires the design of an appropriate anchor size to improve the accuracy of the target location. In order to make the anchor size closer to the ground truth box size of PCB components and thus improve the detection accuracy of small targets, this paper uses the bisecting K-means algorithm to derive a better anchor template from the PCB component training set. Firstly, the algorithm performs cluster analysis on the data to obtain a set of cluster center points. Then, the corresponding anchor dimensions are calculated based on these cluster center points. Lastly, these anchor dimensions are applied to the Faster R-CNN model for both training and testing. Detailed steps are outlined as follows:

1) Firstly, w and h of each component bbox (bounding box) are obtained from the PCB data set, and the mean value of all bboxes is selected as the centroid. All bboxes are considered as a cluster. K-means clustering is performed on each cluster, and K, the number of clusters, is set to derive the initial clustering prior box, that is, (W_i, H_i), $i \in \{1, 2, 3, \ldots, k\}$, where k is the total number of PCB components.

2) The IOU value between each ground truth bbox and the prior bbox is calculated through (3).

$$\text{IOU} = |\text{A} \cap \text{B}| \Big/ |\text{A} \cup \text{B}| \tag{3}$$

A denotes the ground truth bbox position of the PCB defect point; B denotes the prior bbox position.

3) Calculate the distance of each bbox from each cluster by (4).

$$\text{DIS} = 1 - \text{IOU} \tag{4}$$

According to DIS, the ground truth bbox is classified as the closest initial prior bbox to form one cluster, and then the family is divided into two families.

4) Calculate the average value of all classes through (5).

$$W^* = \sum W_i \Big/ n, H^* = \sum H_i \Big/ n \tag{5}$$

where W_i and H_i are the width and height of each bbox, respectively, W^* and H^* are the average of all points in the cluster.

5) Select clusters that meet the criteria and can be decomposed. Considering the number of elements in the cluster and the clustering cost (sum of squares of errors, SSE), the cluster with the largest SSE is selected to be divided into two. Repeat step 3) $\sim$ 5) to updated until it is divided into K clusters and the anchor box is obtained. SSE is calculated as shown in (6) and (7).

$$\text{SSE}_W = \sum_1^n a_i \left(W_i - W^*\right)^2 \tag{6}$$

$$\text{SSE}_H = \sum_1^n a_i \left(H_i - H^*\right)^2 \tag{7}$$

where a_i denotes the weight, W_i and H_i are the width and height of each anchor box, respectively, W^* and H^* are the average values of all points in the cluster.

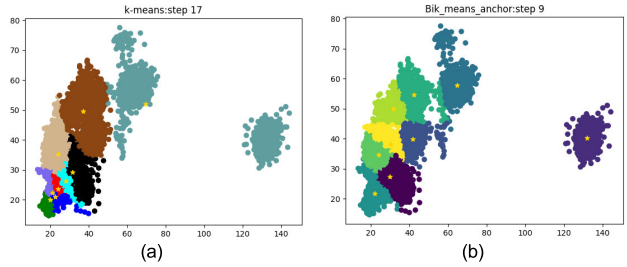

FIGURE 6. Size of the anchor box obtained by the clustering algorithm. (a) K-means cluster results; (b) Dichotomous K-means cluster results.

After above steps, the anchor template dimension matching the PCB components are obtained. The anchor dimensions obtained using the K-means clustering algorithm and the dichotomous K-means clustering algorithm are shown in Figure 6.

The nine anchor sizes obtained by K-means clustering are: (19, 19), (21, 20), (21, 22), (24, 23), (28, 26), (24, 35), (31, 29), (37, 49), and (69, 51). And the use of dichotomous K-means clustering yielded nine anchor sizes: (21, 21), (30, 27), (24, 34), (30, 38), (31, 49), (41, 39), (42, 54), (64, 57), and (132, 40). As can be seen in Figure 6, the dichotomous K-means algorithm outperforms the K-means algorithm on the PCB dataset used in this paper. This superiority can be attributed to several factors: 1) the dichotomous K-means algorithm is much easier to obtain the global optimal solution than the K-means algorithm; 2) the SSE algorithm makes the clustering error minimal each time; and 3) the dichotomous K-means algorithm is not affected by the initial random particle.

IV. EXPERIMENTATION AND ANALYSIS

A. PCB DATASETS

The PCB dataset used in this paper was collected during the actual production process at an electronics factory, and the acquisition equipment is shown in Figure 7. Taking into account factors such as detection accuracy, detection speed, economic cost, data volume, Hikvision MV-CE200-11UC face array camera was selected. In order to clearly capture high-definition images of PCB components with a pixel size of 3629 × 3585, the lens should be:

1) High resolution: The lens must possess a high degree of clarity to produce more detailed images.

2) Uniform picture clarity: The lens must maintain consistent clarity across different areas to prevent image distortion.

3) Ultra-low distortion: The lens should have a low distortion rate to accurately reproduce the shape and size of objects.

4) High peripheral brightness ratio: the lens must maintain a high brightness contrast in the image edge and the center of the region.

According to the inspection environment and camera parameters, Hikvision's 25-megapixel MVL-KF1224M-25MP fixed-focus industrial inspection lens has been chosen.

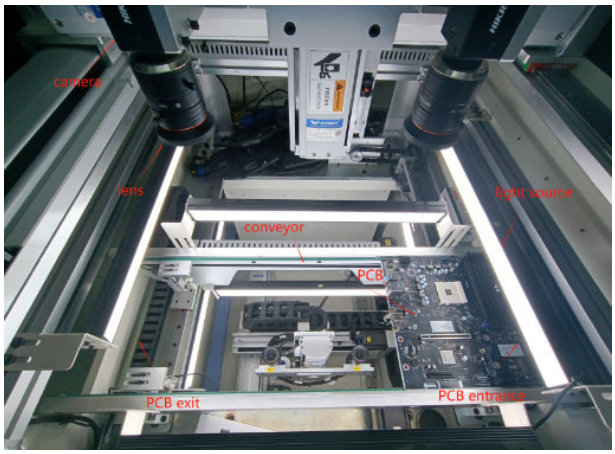

FIGURE 7. PCB dataset acquisition hardware devices.

In terms of light source, we require a unit with flexible angle of light irradiation, compact size, capability of multi-surface installation, stable light emission, high brightness, and high-quality LED lights. In this research, Hikvision MV-LLDS-H-400-40-W is chosen as light source, as it fulfills the stated criteria.

In our research, a PCB image acquisition hardware device using forward illumination is designed. Given the reflectivity of the PCB board, it is crucial to provide uniform light distribution around the board. Therefore, multiple strip light sources are used to illuminate the PCB board, enhancing the imaging quality of the PCB image.

The dataset comprises 1990 high-definition PCB component images with the size of 3629 × 3585. Through a range of augmentation processing, including blurring, mirroring, cropping, and rotating, the dataset has been extended to 3000 images. The dataset contains 12 types of soldered PCB components, which are divided into training set and test set in a 9:1 ratio.

In this dataset, there are 1550 PCB component images with defect and 1450 without defect. The location and category information of PCB components are marked by labelImg. The number and the sample of each category are shown in Table 2 and Figure 8.

B. ENVIRONMENT BUILD

The experimental platform is equipped with Intel (R) Xeon (R) Platinum 8358P CPU @ 2.60GHz, and NVIDIA RTX A5000 (24GB). The software is Python 3.9.7, CUDA 11.6, and Pytorch 1.11.

C. EVALUATION INDICATORS

In order to evaluate the performance of the model in detecting PCB components, Precision (P), Recall (R), Average Precision (AP) for a single category, Mean Average Precision (mAP) for the model, F1, and mF1 are introduced.

$$P = \frac{TP}{TP + FP} \tag{8}$$

TABLE 2. PCB dataset.

PCB components	Number of components (pcs)
Cap_blue_black	4375
Cap_cross	3411
Jumper04p	1568
HDD	2988
Speaker	1960
Power24p	1990
Bat	1950
Power04p	2440
2USB	1850
Jumper10p	2130
Power08p	1980
Rj45+2USB	1894

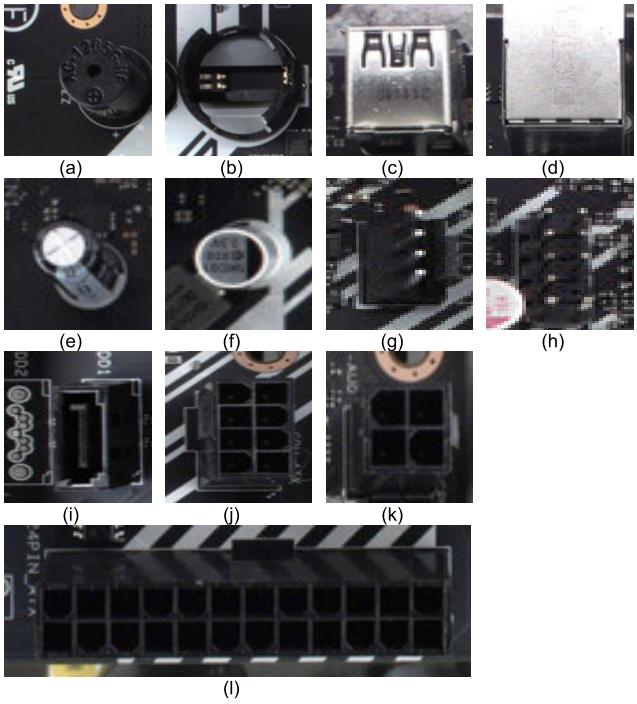

FIGURE 8. PCB component defect detection categories. (a) Speaker; (b) Bat; (c) 2USB; (d) Rj45+2USB; (e) Cap_cross; (f) Cap_blue_black; (g) Jumper04p; (h) Jumper10p; (i) HDD; (j) Power08p; (k) Power04p; (l) Power24.

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{9}$$

$$\mathrm{AP} = \int_0^1 \mathrm{P}(r)\,\mathrm{d}r \tag{10}$$

$$\mathrm{mAP} = \frac{\sum_{i=1}^{k} \mathrm{AP}_i}{k} \tag{11}$$

$$\mathrm{F1} = \frac{2\mathrm{PR}}{\mathrm{P} + \mathrm{R}} \tag{12}$$

$$\mathrm{mF1} = \frac{\sum_{i=1}^{k} \mathrm{F1}_i}{k} \tag{13}$$

P is the proportion of correct prediction frames among the currently traversed prediction frames, high P indicates good performance. R is the proportion of currently recognized labels among all labels, the higher the recall, the better the model performance. mAP is the average model accuracy, the higher the value, the better the model performance, where k is the total number of categories in the recognition target. F1 is the average sum of accuracy and recall, the higher the value, the better the model performance. mF1 is the mean of F1 of all categories. And, TP: True Positive, FP: False Positive, TN: True Negative, FN: False Negative.

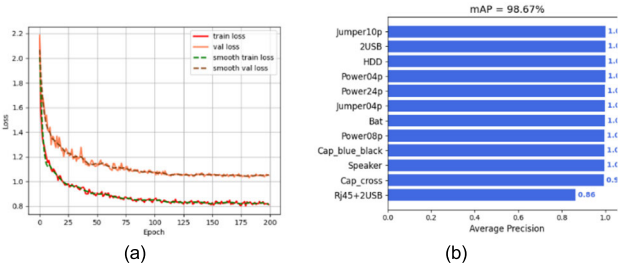

FIGURE 9. Model training loss and model mAP accuracy. (a) Model training loss; (b) mAP accuracy.

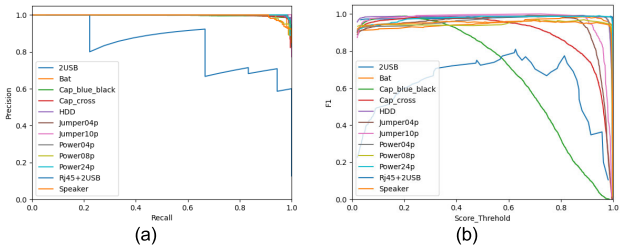

FIGURE 10. Improved model evaluation P-R curve and F1 curve for each defect category. (a) P-R curve; (b) F1 curve.

TABLE 3. Comparison of optimization effects with the introduction of different improvement modules.

VGG 16	ResNet 50	ResNet 101	FPN	K-means	Dichotomous K-means	ECA-Net	Soft-NMS	mAP
√								83.47%
	√							90.34%
		√						89.28%
	√		√					96.26%
	√		√	√				95.78%
	√		√		√			97.45%
	√		√		√	√		98.67%
	√		√		√	√	√	98.67%

D. MODEL TRAINING AND PERFORMANCE EVALUATION

The experimental parameters of the improved algorithm model in this paper are set as follows: the maximum number of iterations is 200, each epoch will be trained iteratively using the training set and validation set, half precision training, batch size = 24, Adam optimizer is used, the initial learning rate is 0.0001, momentum = 0.9, weight decay = 0, the learning rate decreases by step, and the model is evaluated using the coco toolbox. Real-time model training loss curve

TABLE 4. AP values of different methods for detecting defects in different PCB components.

	MSF-ECANet (proposed)	RAR-SSD [26]	YOLOv3 [12]	YOLOv4 [29]	YOLOv5-x [30]	YOLOx-s [31]	Faster R-CNN [14]
2USB	1.00	1.00	1.00	0.98	1.00	1.00	0.99
Bat	1.00	0.98	0.99	0.87	0.99	0.99	0.99
Cap_blue_black	1.00	1.00	1.00	0.94	1.00	1.00	**0.69**
Cap_cross	0.99	0.99	0.99	0.81	1.00	0.99	0.91
HDD	1.00	1.00	1.00	0.89	1.00	0.99	0.97
Jumper04p	1.00	0.99	0.97	0.77	1.00	0.97	0.83
Jumper10p	1.00	1.00	0.60	0.00	0.88	0.97	0.89
Power04p	1.00	0.98	0.98	0.00	0.99	0.99	**0.83**
Power08p	1.00	0.99	1.00	0.26	0.99	0.99	0.98
Power24p	1.00	0.99	0.99	0.99	1.00	1.00	1.00
Rj45+2USB	0.86	0.78	0.60	0.00	0.81	0.31	0.79
Speaker	1.00	1.00	0.99	0.00	0.99	0.97	0.99

TABLE 5. Comparison of mAP, mF1, model size, params, FLOPs (floating-point operations per second) among different methods.

	MSF-ECANet (proposed)	RAR-SSD [26]	YOLOv3 [12]	YOLOv4 [29]	YOLOv5-x [30]	YOLOx-s [31]	Faster R-CNN [14]
mAP@0.5	**98.67%**	**97.26%**	92.61%	**54.24%**	**97.05%**	92.93%	90.34%
mF1	**0.94**	0.91	0.82	0.44	0.91	0.80	0.84
Model size (MB)	472	96.2	235	244	333	**35.8**	108
Params (M)	26.85	25.22	61.58	64.00	87.32	9.32	8.54
FLOPs (G)	**808.58**	**1949.93**	1243.59	480.26	**1744.25**	180.65	367.50

is shown in Figure 9(a); mAP values of the model are shown in Figure 9(b).

As seen from Figure 9(a), the model converged. However, due to the small sample size of the dataset, the loss differs between training and validating. Once the training fished, the models generated in each epoch are traversed, and the recognition accuracy is calculated to determine the best model. The best model is used to evaluate the performance on the dataset consisting of 1263 untrained PCB component images. The P-R curve and the F1 value curve for each PCB component detection category are plotted in Figure 10(a) and Figure 10(b), respectively.

It can be seen from Figs 9 and 10 that the recognition rate of the proposed model for Rj45+2USB components is relatively low. This is primarily attributed to the intense reflection of light from the metal surface of Rj45+2USB components. Although preprocessing is introduced, little improvement achieved.

In order to evaluate the impact on the defect detection performance of PCB components of each improvement module added in the MSF-ECANet model in this paper, experiments were conducted to quantitatively analyze the experimental effect of each improved module based on a test sample dataset. As shown in Table 3, the detection accuracy of the traditional Faster R-CNN algorithm with VGG16 is 83.47% on the PCB dataset. In the proposed model, by using ResNet50 to extracts the feature of PCB components, the accuracy is improved by 6.87% based on the Faster R-CNN algorithm. When the single scale feature extraction is changed to FPN, the accuracy is improved to 96.26%. When the dichotomous K-means algorithm is used in the network model to generate the anchor template, the model accuracy is improved by 1.19%. With the introduction of the ECA-Net module, the overall accuracy of the MSF-ECA-Net model is improved by 15.2% on the basis of the original Faster R-CNN. Through the above analysis, the improved network in this paper has a great detection accuracy improvement effect than the traditional Faster R-CNN algorithm, thus making it more suitable for PCB practical industrial production.

E. COMPARISON OF EXPERIMENTAL RESULTS

To further validate the detection capability of the improved algorithm, the seven algorithms, namely the MSF-ECANet model in this paper, SSD, YOLOv3, YOLOv4, YOLOv5-x, YOLOx-s, and Faster R-CNN, were evaluated on common test dataset consisting of 1263 PCB images. The performance is evaluated by AP value, mAP value, mF1 value, network model size, number of parameters and computation (FLOPs, floating-point operations per second). Corresponding experiments results are shown in Table 4 and 5.

From the experimental data in Table 4, the AP values of the Faster R-CNN network are only 0.69 and 0.83 on

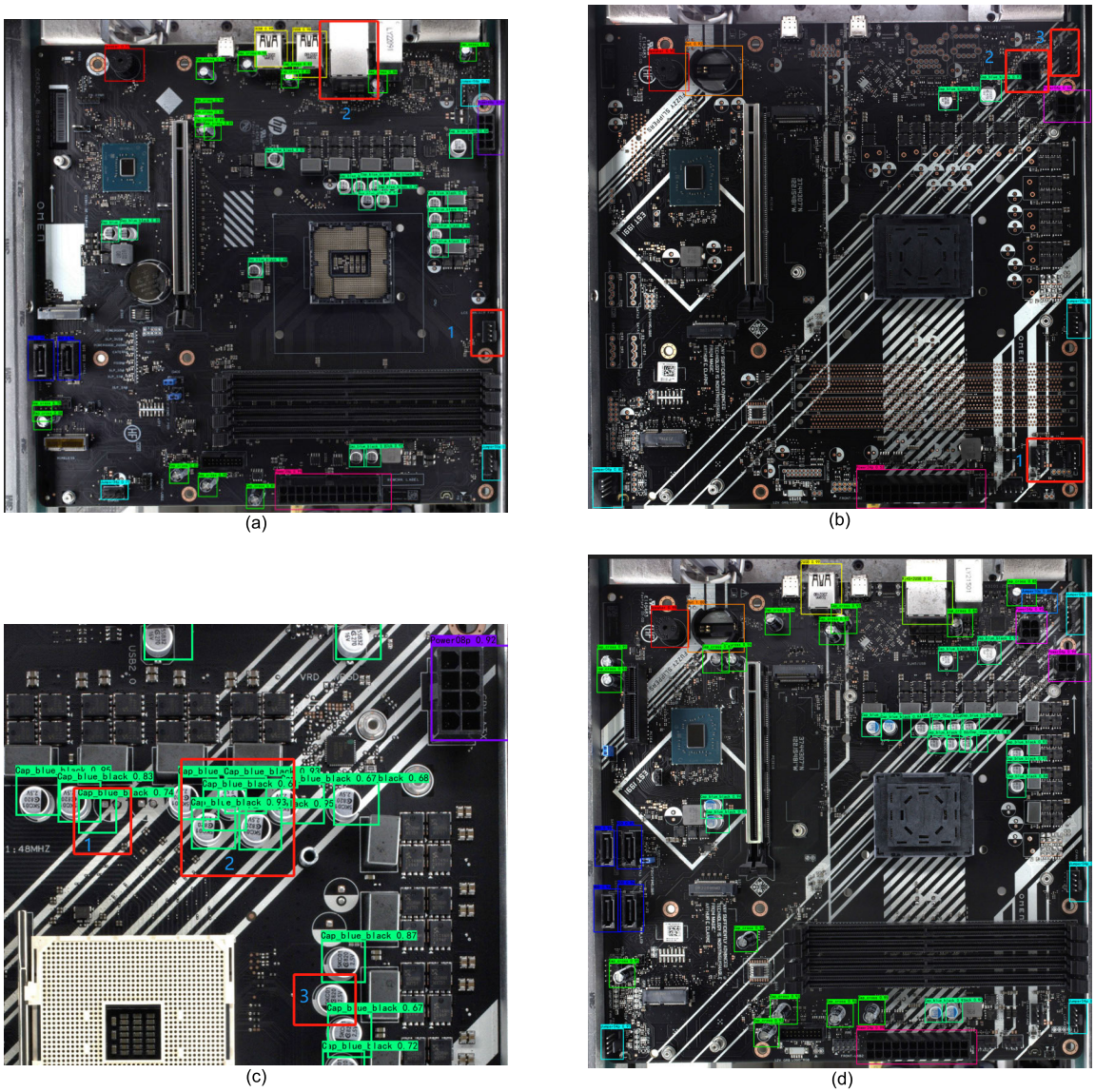

FIGURE 11. Comparison of actual PCB detection between original Faster R-CNN and the proposed MSF-ECANet. (a)Faster R-CNN: miss detections of Rj45+2USB and Jumper04p; (b)Faster R-CNN: miss detections Power04p and Jumper04p; (c) Faster R-CNN: duplicate, miss, and false detection of Cap_blue_black; (d)MSF-ECANet model.

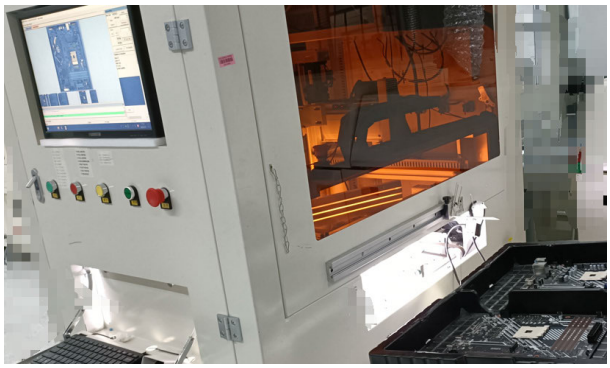

FIGURE 12. Application of MSF-ECANet model in industry.

tiny components (Cap_blue_black) and components with highly overlapped backgrounds (Jumper04p), whereas the AP values of the proposed MSF-ECANet on Cap_blue_black and Jumper04p reached 1.

As shown in Table 5, Both the SSD model and the YOLOv5 model are characterized as lightweight networks, with mAP values of 97.26% and 97.05%, respectively. They are less effective in extracting features of PCB component compared to the improved model presented in this paper. Furthermore, their computational complexity is twice that of improved model. Among all the models, YOLOv4 model exhibited the poorest performance, with a mAP of 54.24%.

So, in summary, despite being the largest among the compared methods, the MSF-ECANet model exhibits the highest mAP value, indicating superior detection accuracy. Although the number of model parameters and the computational effort fall within the middle range of the compared methods, and the accuracy of the model in identifying each

type of PCB components is higher than that of the compared methods.

Figure 11(a)∼(c) are the detection of PCB component using the original Faster R-CNN model. In Figure 11(a), mark 1 indicates a miss detection of Jumper04p, and mark 2 indicates a miss detection of Rj45+2USB. In Figure 11 (b), mark 1 and 3 indicate miss detections of Jumper04p, and mark 2 indicates miss detection of Power04p. In Figure 11 (c), mark 1 indicates false detection, mark 2 indicates duplicate detection, and mark 3 indicates miss detection of Cap_blue_black. Figure 11(d) shows the example of the proposed MSF-ECANet model for detecting defects in PCB components, and it can be clearly seen that the detection capability of the improved model is significantly improved. The practical application of the model in industry is shown in Figure 12.

V. CONCLUSION

This paper investigated two major problems in PCB defect detection: the insufficient detection capability of densely packed small targets and the difficulty in segmenting due to the high similarity between foreground and background information. To address these problems, we propose a method named MSF-ECANet. In the MSF-ECANet method, the Faster R-CNN was used as the basic model. To solve the vanishing gradient problem and increase the receptive field of the model, ResNet and FPN was fused. This fusion enables the model to extract finer PCB details in the deep network. In addition, the ECA-Net model was used, which effectively segment the foreground and background features of PCB components, and weight the foreground more heavily to improve recognition accuracy. Experimental results show that the MSF-ECANet method outperforms algorithms like YOLOv3, YOLOv5, and RAR-SSD in terms of accuracy and efficiency. The MSF-ECANet model can be applied to the PCB production process in electronic factories for PCB defect detection.

Although the MSF-ECANet model can be applied to PCB component detection, metal component detection poses specific challenges. Specifically, the strong reflection on the surface of metal components and the noise problem in image acquisition are hard to deal with. In order to solve these problems, in future research, following strategies could be considered: 1) Light field analysis method could be used to arrange the light source accurately; 2) Image preprocessing technology, such as high dynamic range (HDR) imaging technology, local contrast enhancement, and methods from other research fields like intelligent remote sensing [32], [33], these method may improve the image quality degraded by reflection; 3) Considering the difference between training loss and validation loss, regularization methods, such as L2 weight decay regularization, might be beneficial.

REFERENCES

[1] Z. Zhong and Z. Ma, "A novel defect detection algorithm for flexible integrated circuit package substrates," *IEEE Trans. Ind. Electron.*, vol. 69, no. 2, pp. 2117–2126, Feb. 2022.

[2] A. M. Darwish and A. K. Jain, "A rule based approach for visual pattern inspection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 1, pp. 56–68, Jan. 1988.

[3] T. Vafeiadis, N. Dimitriou, D. Ioannidis, T. Wotherspoon, G. Tinker, and D. Tzovaras, "A framework for inspection of dies attachment on PCB utilizing machine learning techniques," *J. Manage. Anal.*, vol. 5, no. 2, pp. 81–94, Apr. 2018.

[4] 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.

[5] X. Y. Zhao, Y. T. Zhou, F. He, S. Wang, and Z. W. Zhang, "Hierarchical extraction matching printed circuit board components defect detection," *Instrum. Technique Sensor*, vol. 10, no. 8, pp. 84–89, Aug. 2018.

[6] B. Ren and L. Cheng, "Analysis and optimal design of illuminator of AOI vision system," *Control Autom.*, vol. 25, no. 27, pp. 42–44, 2009.

[7] T. Hu, B. P. Guo, and X. Guo, "Contour feature based on image registration," *Opto-Electron. Eng.*, vol. 36, no. 11, pp. 118–122, 2009.

[8] J. Yao, Y. T. Ye, and J. Zhang, "Research of parsing Gerber file in PCB automatic optical inspection," *Comput. Eng. Des.*, vol. 33, no. 6, pp. 2481–2485, 2012.

[9] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, 2004.

[10] Y. Du, F. Yang, and X. Y. Wang, "Research on edge detection algorithm of color PCB image," *Video Appl. Project*, vol. 35, no. 13, pp. 113–115, 2011.

[11] N. S. Qiao, L. Deng, Y. B. Ceng, and J. C. Zou, "Study of noisy and darker PCB photoelectricity image edge detection," *J. Optoelectron., Laser*, vol. 24, no. 4, pp. 740–745, Apr. 2013.

[12] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1–6.

[13] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "SSD: Single shot MultiBox detector," presented at the 14th Eur. Conf. Comput. Vis., Amsterdam, The Netherlands, Oct. 2016.

[14] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[15] Y.-S. Deng, A.-C. Luo, and M.-J. Dai, "Building an automatic defect verification system using deep neural network for PCB defect classification," in *Proc. 4th Int. Conf. Frontiers Signal Process. (ICFSP)*, Sep. 2018, pp. 145–149.

[16] C. Zhang, W. Shi, X. Li, H. Zhang, and H. Liu, "Improved bare PCB defect detection approach based on deep feature learning," *J. Eng.*, vol. 2018, no. 16, pp. 1415–1420, Nov. 2018.

[17] Y.-L. Lin, Y.-M. Chiang, and H.-C. Hsu, "Capacitor detection in PCB using YOLO algorithm," in *Proc. Int. Conf. Syst. Sci. Eng. (ICSSE)*, New Taipei City, Taiwan, Jun. 2018, pp. 1–4.

[18] G. Z. He and L. Yu, "PCB defect detection based on convolutional neural network," *J. Graph.*, vol. 43, no. 1, pp. 21–27, 2022.

[19] Z. C. Sun, B. Wang, and X. L. Zhang, "PCB defect detection based on deformable residual convolution and scalable feature pyramid algorithm," *Telecommun. Eng.*, vol. 63, no. 6, pp. 798–805, Jun. 2023.

[20] Z. Zhang, M. Zhou, H. Wan, M. Li, G. Li, and D. Han, "IDD-Net: Industrial defect detection method based on deep-learning," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, Art. no. 106390.

[21] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 580–587.

[22] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1440–1448.

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

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[25] H. Y. Zhu, Z. P. Li, Y. Zhao, X. H. Luo, X. J. Cheng, and X. W. Yang, "PCB defect detection algorithm based on multi-scale fusion and deformable convolution," *Comput. Eng. Des.*, vol. 43, no. 8, pp. 2188–2196, Aug. 2022.

[26] W. Jiang, T. Li, S. Zhang, W. Chen, and J. Yang, "PCB defects target detection combining multi-scale and attention mechanism," *Eng. Appl. Artif. Intell.*, vol. 123, Aug. 2023, Art. no. 106359.
[27] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 936–944.
[28] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, "ECA-Net: Efficient channel attention for deep convolutional neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 11531–11539.
[29] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.
[30] *YOLOv5*, Ultralytics, Frederick, MD, USA, 2021.
[31] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO series in 2021," 2021, *arXiv:2107.08430*.
[32] D. Hong, B. Zhang, X. Li, Y. Li, C. Li, J. Yao, N. Yokoya, H. Li, P. Ghamisi, X. Jia, A. Plaza, P. Gamba, J. A. Benediktsson, and J. Chanussot, "SpectralGPT: Spectral remote sensing foundation model," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Apr. 3, 2024, doi: 10.1109/TPAMI.2024.3362475.
[33] X. Wu, D. Hong, and J. Chanussot, "UIU-Net: U-Net in U-Net for infrared small object detection," *IEEE Trans. Image Process.*, vol. 32, pp. 364–376, 2023.

WENBIN CHEN received the Ph.D. degree in information and communication engineering from the North University of China, in 2019. She joined the School of Electrical Engineering, Chongqing University of Science and Technology, in 2020. Her research interests include medical image processing, PCB defect detection, and deep learning.

HONGCHAO ZHAO received the B.S. and M.S. degrees from the School of Electrical Engineering, Chongqing University of Science and Technology, Chongqing, China, in 2019 and 2023, respectively. His current research interests include computer vision, image processing, and machine learning.

ZHENG WANG received the B.S. and M.S. degrees from the School of Electrical Engineering, Chongqing University of Science and Technology, Chongqing, China, in 2019 and 2022, respectively. He is currently pursuing the Ph.D. degree with the Graduate School of Engineering, Mie University, Japan. His current research interests include machine learning, machine vision, and deep learning.