Neighborhood Correlation Enhancement Network for PCB Defect Classification

Wenkai Huang, *Member, IEEE*, Jiafu Wen, Weiming Gan, Lingkai Hu, and Bingjun Luo

***Abstract*— In recent years, deep-learning has been gradually applied to the detection and classification of true and pseudo defects in printed circuit boards (PCBs). To judge the authenticity of PCB defects, it is necessary not only to combine the shape characteristics of the defects themselves, but also to add the relationship information between the surrounding environment and defects. However, it is difficult for general classification methods to extract such features. In this article, we propose the neighborhood correlation enhancement network (NCE-Net), which effectively uses defect and surrounding relationship information to accurately distinguish defect authenticity. This network has a relevance residual block (RRB), which is used to establish the correlation between defects and their surroundings, including the location enhancement and locate (LEL) module and relevance convolution (RC), which are respectively used to enhance the effective geographical information and extract the relationship between features of different positions. It also utilizes a small squeeze residual block (SRB) to classify pseudo defects more quickly and efficiently in industrial applications. In addition, to achieve the lowest pseudo-defect detection error rate, we created a unique multi-network specific integration (SI) method for use with NCE-Net. The experimental results show that our proposed network can be trained on a PCB defect classification dataset (PCB-2-DET) for higher efficiency and more significant PCB defect detection. Additionally, the model's identification accuracy can be further improved through our unique SI method.**

***Index Terms*— Attention mechanism, neighborhood correlation, printed circuit board (PCB) welding defect detection, relevance convolution (RC).**

I. Introduction

IN THE production and manufacturing process of the printed circuit board (PCB), it is easy to have defects that will affect its performance, such as open circuit, short circuit, notch, residual copper, etc., so defect detection is a very important link to ensure the good performance of the PCB. Specifically, in PCB defect detection, automatic optical inspection (AOI) is used to carry out strict pixel-level image contrasts via image recognition [1], [2], and template-based matching [3], which are simple and easy to operate. However, these traditional methods are very sensitive to light changes, geometric invariance, and inaccurate registration, among other factors [4]. As a result, it is difficult to accurately identify defects with complex shape changes, resulting in a high pseudo-alarm rate. Therefore, after AOI defect detection, a manual reinspection link is usually set up to distinguish whether the defect is true or pseudo. Even so, manual detection is inefficient and costly, making it difficult to meet PCB batch production requirements.

In recent years, computer vision (CV) technology has developed rapidly, with more and more CV algorithms applied to detect PCB defects in industrial production. In the beginning, traditional image processing technologies, such as template matching, corner detection, and scale-invariant feature transform feature matching, were used. However, these methods can only manage graphics that are fixed and simple, so it is still difficult to distinguish defects mixed with complex backgrounds or that are rich in abstract rules. With its remarkable progress in deep learning, though, image recognition has begun to replace traditional methods and is now often applied to various defect detection tasks in industrial manufacturing. There are three main types of detection tasks, the first of which is directly classifying defects. This task's main representative methods involve the depth convolution neural network [5], depth residual neural network (ResNet) [6], ResNet [7], [8], [9] with an attention mechanism, and measurement learning [10] for the detection of small defect samples. Recently, Fang et al. [11] proposed a fabric structure defect detection method that combines tactile sensors and convolutional neural networks (CNN) for defect classification. The second task is target detection, which involves analyzing the "what" and "where" of defects. Its framework structure includes Backbone, Neck, and Head, and it is divided into one-stage object detection and two-stage detection algorithms. The one-stage object detection method can directly obtain the classification probability and location coordinate value of the object, which can achieve faster detection speed; recently Shengqi et al. [35] proposed an improved YOLOv5-based target detection method to achieve the detection of ceramic ring defects [36]. Efficient PCB detection can be realized through a fast defect detection network based on mobilenet-Yolo-fast. The two-stage object detection method first extracts the region proposals of the object and then classifies them, which can have higher accuracy. The two-stage method, such as that described

Manuscript received 7 November 2022; revised 29 December 2022; accepted 19 January 2023. Date of publication 9 February 2023; date of current version 20 February 2023. This work was supported by Guangzhou Youth Science and Technology Education Project under Grant KP2023243. The Associate Editor coordinating the review process was Dr. Chuang Sun. *(Corresponding authors: Wenkai Huang; Bingjun Luo.)*

Wenkai Huang, Jiafu Wen, Weiming Gan, and Lingkai Hu are with the School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China (e-mail: smallkat@gzhu.edu.cn; 1907700024@e.gzhu.edu.cn; 2112107050@e.gzhu.edu.cn; 2112107060@e.gzhu.edu.cn).

Bingjun Luo is with Guangdong Greatsense Intelligent Equipment Company Ltd., Guangzhou 510006, China (e-mail: bingjun.luo@polytechoptics.com).

Digital Object Identifier 10.1109/TIM.2023.3243670

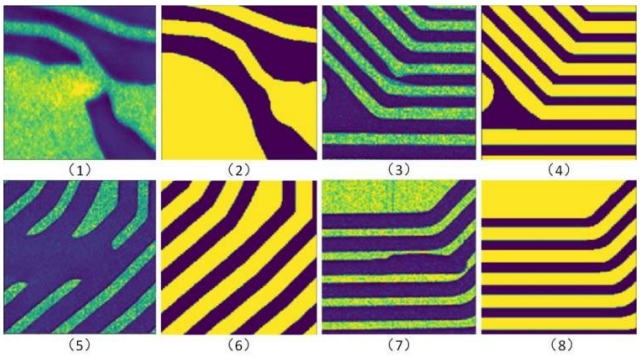

Fig. 1. (1), (3), (5), and (7) are the defect maps and (2), (4), (6), and (8) are the standard maps. The labels of (1), (2), (5), and (6) are true defects, while the labels of (3), (4), (7), and (8) are pseudo defects.

in [12], a two-stage decoupling target detection framework is proposed to detect defects in flexible PCBs [13] recently outlined pooling- balanced-feature pyramid network method to more accurately detect different kinds of defect vision information, and achieve more accurate defect location and classification. The third and final task is example segmentation in which defects are classified pixel by pixel. One typical method for this task is U-Net [14], recently Ling et al. [15] developed a twin semantic defect segmentation network that integrates multi-level feature maps. Two channels input PCB standard maps and defect maps at the same time, which can yield accurate PCB defect segmentation with fewer samples. However, these methods are only applicable to conventional PCB defect detection tasks. In the PCB true and false defect detection task, the recognition of each category is lower, so a detection model that pays more attention to the details of true and pseudo defects is still needed to accurately distinguish them.

The judgment of true and pseudo-PCB defects is equivalent to the judgment of the defects' functional impact. Defect classification can also be called functional defect classification according to whether the defects affect the PCB's original function. Defect classification's key feature is not only identifying the shape of the defects themselves but also understanding the defects' surrounding environment. For instance, as shown in Fig. 1(1), when a notch-style defect appears on a circuit in a PCB, if the notch cuts off the circuit, the defect affects the circuit's continuity, thus affecting its original function. This is considered a functional (true) defect, and the exact defect is an "open circuit." Meanwhile, in Fig. 1(3), if the defect appears on a circuit in the PCB but the original circuit is not cut off, the defect does not affect circuit continuity, so it does not affect its original function. This type of defect is then regarded as nonfunctional (pseudo). Therefore, to achieve high-precision PCB defect detection, it is necessary to have the model consider the relationship between defects and their neighborhood information. Although the existing CNN-based neural network can gradually extract defects in input data through deep operations, it fails to accurately extract the relationship between defects and their surrounding neighborhood backgrounds to achieve high-precision classification. To solve this problem, Bao et al. [16] proposed a triple graph inference network for feature segmentation, as the network's inference can more flexibly extract the relationship between defect and environmental features and achieve fine defect segmentation. EMRA-Net in [17] realized the association between global and local features by fusing global dynamic convolution features, global dynamic multi-scale fusion features, and local pyramid edge features. In [11], the combination of CNN and effective channel attention modules improved channel correlation and strengthened the connection between local features. While all these methods try to build the relationship between features of different objects in the same feature map, they still cannot meet the strong correlation features of the relationship between defects and surrounding neighborhoods needed during PCB defect detection. Accordingly, there remain challenges in identifying true and pseudo-PCB defects.

To better extract the relationship between defects' relative position and their surrounding background environments and thus enhance defect identification, we propose a neighborhood correlation enhancement network (NCE-Net) that features a squeeze residual block (SRB) and relevance residual block (RRB) based on ResNet network architecture. Among them, ResNet is proposed by Kaiming et al. [37]. The network is basically composed of a feature mapping layer, multiple residual blocks, and a full connection layer. The residual block is composed of multiple cascaded convolution layers and shortcut connections. After the output values of the two are accumulated, the output of the residual block is obtained through the ReLU activation layer. Multiple residual blocks can be connected in series to achieve a deeper network, which to some extent solves the "degradation" problem of deep neural networks.

Specifically, SRB improves the network operation efficiency, while RRB uses location enhancement and locate (LEL) and relevance convolution (RC) to extract effective relationship information. In detail, the LEL module extracts a multi-view map (MVP) and location map (LM) for input data through a multi-view operation. We further use the RC, input MVP and LM to it, associate each point in the graph with the effective information around it, and then output it. Ultimately, our model, the NCE-Net, realizes high-precision defect detection by combining defects with their surrounding background. Compared with previous studies, this network can better detect true and pseudo defects. To further improve the detection of pseudo defects, we also propose a model-specific integration (SI) method that uses multiple different models to judge defects and make decisions to achieve a high classification and detection rate of true and pseudo defects at the small cost of filtering out fuzzy data. Our contributions are summarized as follows:

1) We propose a novel NCE-Net, which is different from the conventional classification network in that it can achieve more accurate classification of real and pseudo-PCB defects by extracting the relationship between defects and surrounding, and also reducing the number of network parameters.

2) Considering that the relationship information of the surrounding areas of defects has a great impact on the authenticity judgment of defects, we propose the RRB module,

including the LEL used to enhance the effective geographical information of the surrounding areas and the RC used to extract the relationship between different location features.

3) Considering the need for a more efficient detection model in actual detection, we propose the SRB module, which can identify small defects with high accuracy with as few parameters as possible.

4) In order to further improve the classification accuracy of true and pseudo defects, we propose a model SI method, which can greatly improve the detection of false defects at the cost of filtering a small amount of fuzzy data.

The rest of this article is organized as follows. Section II reviews the related works on defect detection and attention mechanism. In Section III, we present the PCB-2-DET in detail. Section IV describes the proposed method in detail. Section V gives some experiments and a discussion of the proposed method. Finally, a conclusion is made in Section VI.

II. Related Works

A. Defect Detection

Defect classification mainly starts with traditional CV image processing. In these methods, defects are generally found by aligning images or searching feature points. For example, [32] described a template-matching method to find defects via template subtraction and positioning. Wang et al. [33] developed an AOI system based on machine vision that integrates a variety of hardware and software to quickly check the missing, porous, and wrong positions of holes in PCBs. Jiang et al. [1] proposed an AOI algorithm for digital image processing that can efficiently distinguish the differences between PCB and standard drawings in a variety of ways, as well as quickly and accurately find defects. Additionally, [18] outlined a method combining feature extraction, feature matching, and graph-to-graph subtraction to classify four kinds of PCB defects.

However, the flexibility of traditional CV methods is low, and it is difficult to identify defects with complex abstract rules. Therefore, researchers have introduced machine learning algorithms to defect identification methods. For example, Zhu et al. [34] proposed a PCB solder joint quality detection method based on feature extraction and support vector machine classification. Further, with the rise of the depth convolution neural network, more and more defect detection algorithms are integrating depth learning. Notably, [5] designed a defect classification algorithm based on the convolution neural network with an accuracy rate far higher than that of traditional classification methods. Ghosh et al. [6] recommended classifying true and pseudo-PCB points based on the Inception V3 pretraining model. To improve the detection speed of deep learning models in practical engineering applications while ensuring accuracy, Xia et al. [12] proposed a two-stage algorithm combining the structural similarity index and MobileNet to detect surface defects on PCBs. To enhance defect detection when the number of samples is too small, the measurement learning method based on the twin network [15] is usually used to detect defects or segmentations; for instance, [10] combined the cost-sensitive function and twin network to classify true and pseudo-PCB defects.

B. Attention Mechanism

To build a correlation between features, the attention mechanism is usually added to CNN. The attention mechanism allocates weights to suppress or enhance certain feature information to guide a model to pay more attention to areas of input data deemed important. Our work is similar to an attention mechanism that enhances the focus on local and defect centers. In [7], the researchers proposed the squeeze-and-excitation (SE) attention method, which enhances the correlation between channels by creating channel attention between convolutions. Woo et al. [20], Fu et al. [8] performed double attention enhancement on channels and spaces on a feature map. The Convolutional Block Attention Module presented in [20] similarly combines maximum pool and space maximum features to generate an attention map, while DA-Net in [8] uses a series of space and channel matrix operations to enhance the relationship between single point and global features. Wang et al. [21] described a nonlocal neural network to extract long-distance correlations by calculating the response of a location as the weighted sum of all location features. To obtain more effective attention with less weight, Zhang et al. [9] proposed the Pyramid Squeeze Attention module and designed a simpler and more efficient framework through the combination of pyramid compression and the SE module. Similarly, Hou et al. [22] developed coordinated attention, which embeds location information into channel attention to obtain more location-sensitive attention maps.

III. Dataset

In this section, we describe in detail the process of building the PCB-2-DET and the basic information of the PCB-2-DET.

In the AOI link of the PCB production link, we collected the PCB judged to have defects and obtained the actual PCB defect location pictures through camera photography. Then, we extracted the detailed standard map corresponding to the defect map one by one from the standard map corresponding to PCB by locating the defect location, so we got the data map with defects and the corresponding nondefect standard map data. At the same time, we have carefully classified and labeled the defect maps manually. According to whether the defects affect the original function of PCB circuits, we have divided the defect maps into "functional defects" and "nonfunctional defects," commonly known as "true defects" and "pseudo defects." If the appearance of a defect breaks the circuit, causing the original circuit to fail to connect, or the appearance of a defect causes a short circuit in the circuit, affecting the normal continuity of the circuit, etc., which affects the original function of the PCB, it is classified as a "functional defect," and vice versa. We name the processed dataset PCB-2-DET, and the specific example is shown in Fig. 1. Specifically, the dataset has 6400 defect maps in total, 3200 for true defect maps, and 3200 for pseudo-defect maps. Each defect map has a corresponding standard map and a true or pseudo label.

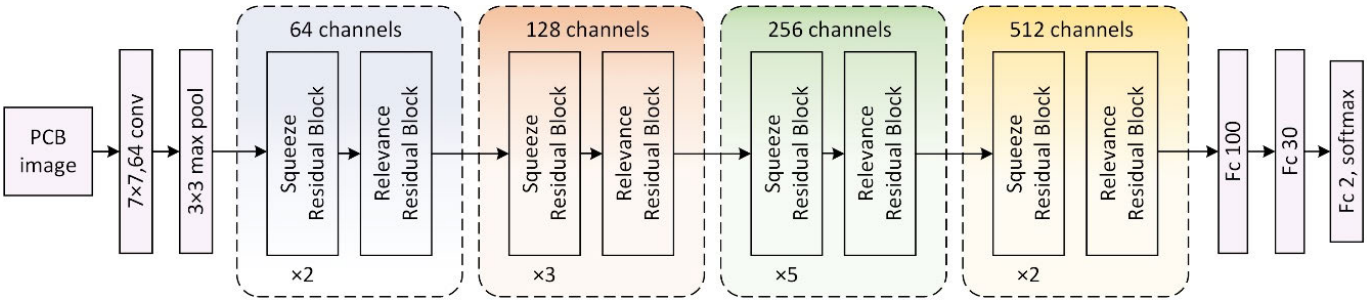

Fig. 2. Structure of the proposed NCF-Net.

IV. Method

A. Network Framework

We developed NCE-Net as a model that effectively and accurately classifies PCB defects. Based on the residual network, we replaced the residual blocks with new SRBs and RRBs so that the model operates with as few parameters as possible and more effectively extracts true and pseudo defects and neighborhood relationships to accurately identify them. Fig. 2 depicts our proposed network, which consists of a feature extractor with four stages of convolution and a classifier with three full connection layers. The residual units in each stage contain multiple SRBs and an RRB. Among them, the RRB is mainly used to build and enhance the relationship between the target defect and its surrounding environment to improve detection accuracy. In turn, the SRB replaces the original residual block with fewer parameters for basic feature extraction to further improve detection efficiency.

Our network first obtains the primary feature map from the input data through convolution and maximum pooling layer, then gradually inputs the data into the four stages of convolution.

Next, the network extracts features through the SRBs to obtain more effective feature maps. Then, the network enhances the relative geographic location information between features in the feature map through RRBs to obtain maps richer in shape, specific features, and correlation information. With this information, it inputs the convolution for the next stage. The number of SRBs in the four stages is 2, 3, 5, and 2, respectively, while there is one RRB at the end of each convolution stage. The final output feature map is flattened into a one-dimensional vector, and the binary classification results are obtained through a fully connected network.

B. Squeeze Residual Block (SRB)

To achieve the highest detection speed in practical industrial applications, we attempted to reduce the number of model parameters as much as possible without affecting accuracy. Accordingly, we introduced SRBs to NCE-Net. As shown in Fig. 3, considering that the effect of convolution feature extraction should be affected as little as possible, we combined the depthwise convolution (DW) and pointwise convolution (PW) from MobileNet [24] with eroding pooling (EP) to replace the original ordinary convolution operation, greatly reducing the number of parameters. First, we used DW

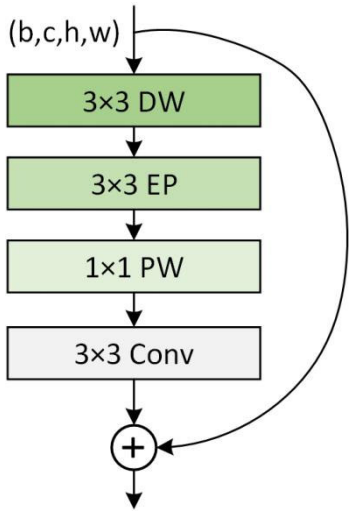

Fig. 3. Squeeze residual block.

to extract a size of [3 × 3]. We filtered the features on the receptive field through the EP of the same receptive field size, which is equivalent to filtering many points with partial features or those with noise features and alleviates redundancy in the features extracted from a complex PCB background. Then, we carried out PW to build the connection between the feature map channels. Finally, after a layer of ordinary convolution, we performed a residual connection and summation to obtain a feature map rich in detailed and effective information.

C. Relation Residual Block (RRB)

The detection of PCB defects has a great relationship with the defects' surrounding environment. For example, if it is judged that a circuit is open, it is necessary to simultaneously judge that the defect's surrounding background environment is a connected circuit and that the defect has disconnected the surrounding circuits. Therefore, to improve the accuracy of PCB defect detection, models must emphasize the relationship between defects and their surroundings. We thus introduced an RRB to enhance the connection between effective features in the input feature graph to guide NCE-Net to pay more attention to the correlation between defects and their neighbors. As shown in Fig. 4, RRB first obtains MVP and LM for the input characteristic diagram through the LEL module. Where, LM size is [b, 18, H, W], b represents the batch

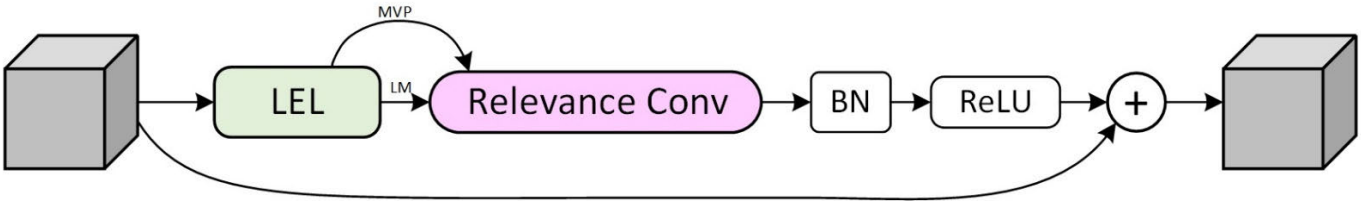

Fig. 4. Relevance residual block.

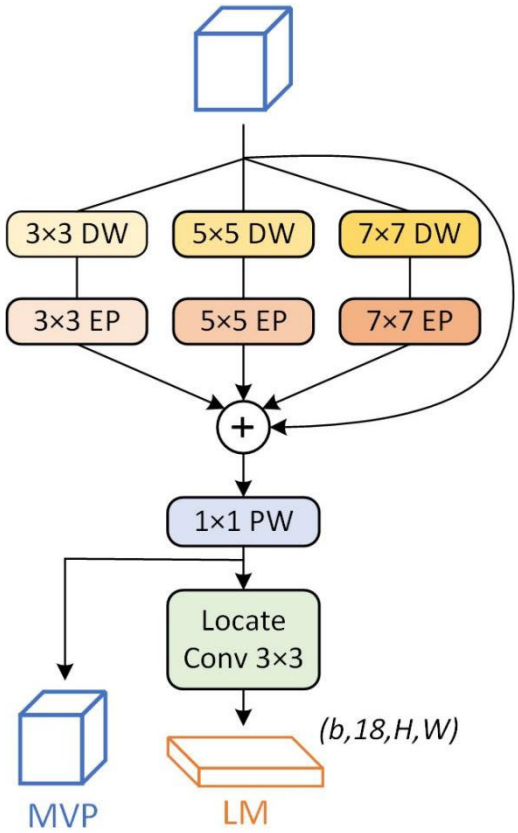

Fig. 5. LEL module.

size, H and W represent the length and width of the input feature map respectively, and 18 dimensions represent nine coordinates (X, Y) corresponding to each point. Then, LM and MVP are input into the relevant convolution at the same time to extract the feature relationship information of any distance between the defect and its surroundings.

Fig. 5 shows the LEL module specifically. Input the obtained feature maps into DW of different receptive fields to obtain feature maps of different visual fields. Then, redundant information can be filtered through the EP of the corresponding size, so that maps of different views can be better aligned and merged together. From there, the MVP can be obtained through a PW calculation. Locale Convolution in Fig. 5 refers to compressing the channel number to 18 through convolution, which can be expressed as follows:

$$\mathrm{LM} = \mathrm{Conv}(\mathrm{inchannels}, 18, 3, 1, 1)(\mathrm{MVP}). \quad (1)$$

The position relationship represented by the values of these 18 channels in LM is not a simple coordinate, but a coordinate offset, which has the same effect as the offset in deformable convolution [25]. It is worth noting that the nine coordinate off-sets of X and Y directions obtained by each spatial point in LM are not limited to the positions of nine adjacent points, but the positions of nine points related to the characteristics of spatial points in the whole map.

Because the shape of defects is changeable and complex, it is difficult for ordinary convolution to fit this flexibility in detail. Therefore, we propose a new RC based on deformable convolution, which not only can make convolution more flexible to fit defects' shape features but also decouple the feature relativity relationship between the defects themselves and their surrounding environment. As shown in Fig. 6, after the LM and MVP are input, RC transforms the LM from shape [b, 18, H, W] to shape [b, 2,3 × H,3 × W] to obtain a two-direction offset map, two channels represent the offset in the x-direction and the offset in the y direction respectively. The offsets are used to index the MVP to obtain a correlation map with the size [b, c, 3 × H, 3 × W]. For a specific method, refer to deformable convolution [25]. The absolute position correlation map is then obtained by splicing the two-direction offset map and correlation map with the size [b, c + 2, 3 × H, 3 × W].

In order to further eliminate the influence of absolute position correlation between features and retain only relative position correlation, we divide the absolute position correlation graph into multiple unit graphs with the size of [1, c, 3, 3] and the step size of 3. We then rotated each unit graph with its own center to 0°, 90°, 180°, and 270°, respectively to obtain four position correlation graphs in different directions. After adding the four graphs to obtain a relative position correlation graph, we at last performed a convolution operation to output the original channel number. The final feature map contained unlimited long- and short-distance peripheral correlation information.

V. Experiments

In this section, we evaluate NCE-Net's performance with the PCB-2-DET dataset and compared it with state-of-the-art (SOTA) methods. We also prove the superiority of SRBs and RRBs in defect detection through a series of comparisons. In addition, through conducting ablation experiments, we show that this article's contributions are mutually compatible.

A. Experimental Settings and Dataset

The experimental environment in this study was run on the Ubuntu system using a 14 core Intel[1] Xeon[1] Gold 6330 CPU

[1]Registered trademark.

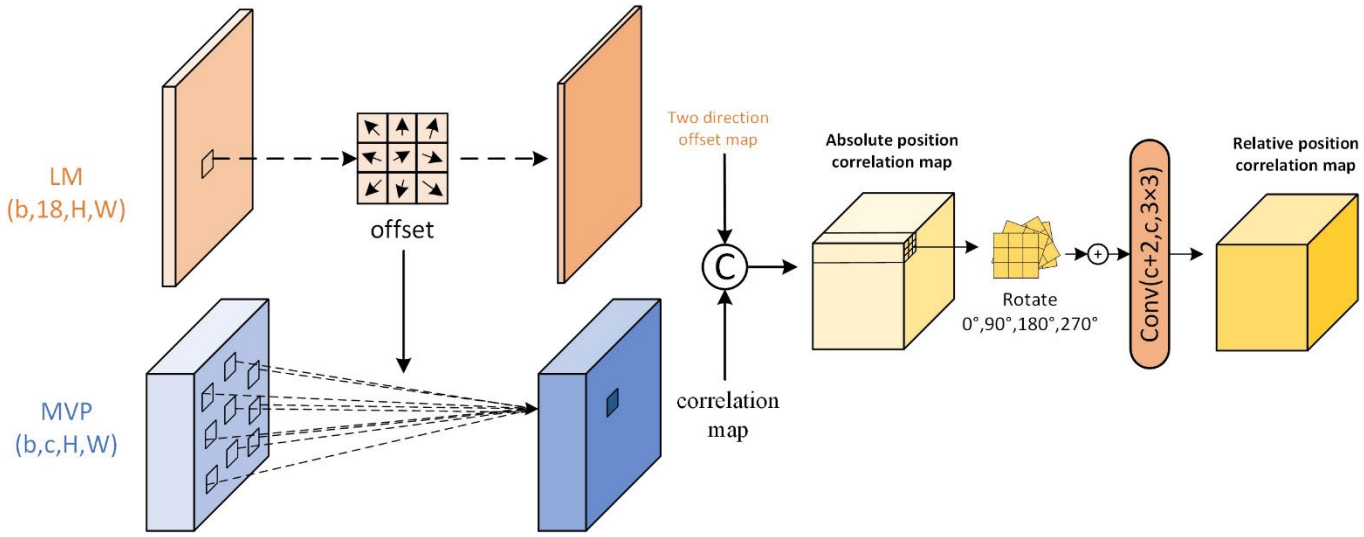

Fig. 6. Relevance convolution.

TABLE I
PCB-2-DET DETAILS

Dataset:PCB-2-DET			Type	Training set	Testing set	Validation set
Image information of dataset	Width	224	Positive	2560	320	320
	Height	224				
	Channels	4	Negative	2560	320	320
	Number	6400				

and a 24-GB RTX3090 video card. The PCB defect dataset used, PCB-2-DET, was proposed by the team for method evaluation. The batch size, epoch number, and learning rate were set to 128, 100, and 0.0001, respectively. All networks used AdamW's gradient descent training method. The momentum, weight decay, and dropout rates were 0.9, 0.0005, and 0.5, respectively.

PCB-2-DET contains real PCB images, corresponding standard board images, and defect authenticity labels. For our study, the images were divided into real- and pseudo-defective board images according to the label. The real defects included circuit open circuit, circuit short circuit, solder mask damage, and other defects that lead to PCB failure. We adjusted the resolution of all images to 224 × 224 to reduce the computational burden. Through the image processing template matching technology, we also aligned the defect map in the defect dataset with the standard map and spliced the map with only a single channel onto the real PCB map. Therefore, all the input data shapes in this experiment were 224 × 224 × 4. We used data enhancement operations such as random rotation and random brightness to normalize the data and improve the model's robustness. As shown in Table I, PCB-2-DET includes 3200 image pairs of real defective plates and their corresponding standard plates, as well as 3200 image pairs of pseudo-defective plates. Among them, 80% were used as the training set, 10% formed the verification set, and 10% acted as the test set.

First, we conducted ablation experiments to compare the difference between the original ResNet and the addition of SRB and RRB in ResNet to prove the effectiveness of the two modules. Second, we compared our model with the most advanced models currently in use. Finally, we test the proposed SI integration method to verify whether it can greatly improve the model's ability to identify true and pseudo defects.

TABLE II
CONFUSION MATRIX

	True defect	Pseudo defect
Predicted as True defect(Positive)	TP	FP
Predicted as Pseudo defect(Negative)	FN	TN

B. Evaluation Metrics

To evaluate the classification of the true and pseudo defects, we divided the prediction results into four groups: true positive (TP), the accurate prediction of true defects as true defects; true negative (TN), the judgment of pseudo defects as pseudo defects; false negative (FN), the judgment of true defects as pseudo defects; and false positive (FP), the judgment of pseudo defects as true defects. We also introduced the confusion matrix, as shown in Table II.

Due to the different requirements for the recognition rate of real defects and pseudo defects in PCB real defect detection tasks, we should not only compare the overall classification accuracy of the model when evaluating the model, but also compare the recall, specificity, precision, negative-precision (NP), and parameter of NCE-Net as evaluation indicators.

We define them as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{2}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{5}$$

$$\text{Negative\_Precision} = \frac{\text{TN}}{\text{FN} + \text{TN}}. \tag{6}$$

Recall rate represents the ratio of predicted real defects to all real defects; specificity indicates the proportion of predicted pseudo defects to all pseudo defects; and the precision ratio represents the ratio between the number of true defects correctly judged and the total number of true defects judged. The negative precision ratio represents the ratio between the number of correctly judged pseudo defects and the total number of judged pseudo defects.

C. Ablation Experiment

1) SRBs: To verify that SRBs can minimize the number of network parameters while maintaining good feature extraction and high accuracy, we will compare the original ResNet with the ResNet with SRB blocks. The convolution layer of the original ResNet used in the experiment has 34 layers, and its total number of residual blocks is 16, respectively: there are three residual blocks processing, 64 channel characteristic residual blocks processing, 128 channel characteristic residual blocks processing, 256 channel characteristic maps, and three residual blocks processing 512 channel characteristic maps. We also compared MobileNetv3 [26] and GhostNet [27], two lightweight networks. Our results show that compared with the original residual network, the residual network with SRB (see Table III) not only greatly reduces the number of parameters, but also has little impact on the network output. Compared with MobileNetV3 and GhostNet, although the number of parameters in our proposed model is greater, its detection accuracy is much higher, thus validating the SRB.

2) RRB: To confirm that RRBs can enhance geographic information and add neighborhood correlations to improve our model's classification accuracy, as shown in Table III, we compared the model's accuracy before and after adding this module. We find that the accuracy and precision of the model in identifying positive and negative samples will be improved no matter whether RRB is directly introduced into the original residual network or RRB is introduced after SRB is introduced. We also compared the difference between ResNet and ResNet with RRB in identifying samples, as shown in Fig. 7. A comparison of the feature maps demonstrates that the RRB module enhances the model's ability to identify defects by building effective neighborhood relationships, filtering out more invalid information, and more accurately identifying defect features. From the last line of the figure, we can also see that the original ResNet does not accurately detect subtle defects in complex environments, while the model with the RRB block can locate the core of the defect and weaken irrelevant background and noise information, proving the RRB module's effectiveness.

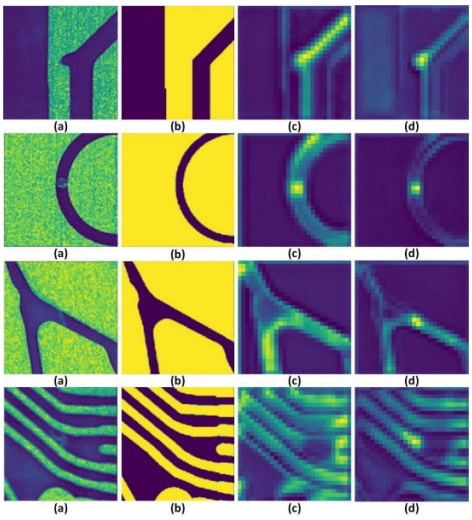

Fig. 7. Comparison of the characteristic map is (a) input image, (b) standard image corresponding to the input, (c) characteristic map during the ResNet calculation, and (d) characteristic map during the ResNet with RRB calculation.

D. Comparison With SOTA Detectors

Using the PCB-2-DET dataset, we further compared NCE-Net with SOTA classification methods, including classic classification methods: sparse and high computing performance network Inception-v4 [28], classic convolutional attention classification network SE-Net [7], and classification network NLNet [21] that can combine long-distance relationships, as well as recent cutting-edge classification methods: GCNet [29], coordinate attention network (CANet) [22], ECA-Net [11] for high-precision classification of fabric defects, and the latest ResNet classification network with attention mechanism: EPSANet [9], MSDSM [30], PCANet [38], and RMCSAM [39]. These methods are mainstream or cutting-edge classification methods. Many of them combine attention mechanism, which also has certain advantages for classifying PCB defects. So we choose these networks to compare with the NCE-Net in this article. Through our experimental comparison, we conclude that ResNet34 can use this data set to achieve high defect recognition accuracy compared with ResNet18, ResNet50, and ResNet101. Therefore, we use ResNet34 as the basic network of all improved attention mechanism models based on ResNet.

As shown in the results in Table IV and the comparison results of test set accuracy of multiple networks in Fig. 8, our method has better classification accuracy and good detection in accuracy, specificity, precision, and negative precision. Although the EPSANet model that achieves the best recall also has good detection capability, its parameter is 287 MB, which is not applicable to actual PCB defect detection. In addition, we also compared the PR curve of the most advanced classification network with the network in this article. As shown in Fig. 9, the PR curve of this model almost covers the PR curve of other models, which shows that the classification network in this article has a better performance.

TABLE III
EFFECT OF DIFFERENT BLOCKS ON NETWORKS

Network	Accuracy	Precision	Recall	NP	Specificity	Param.
ResNet34	87.29%	87.24%	87.60%	87.34%	86.97%	81.484MB
MobileNetV3	76.88%	78.20%	79.69%	75.23%	73.52%	7.020MB
GhostNet	81.46%	83.08%	82.76%	79.55%	79.91%	15.082MB
SRB+ResNet34	87.29%	85.66%	88.96%	88.98%	85.71%	48.473MB
RRB+ResNet34	88.96%	88.34%	89.43%	89.50%	89.32%	82.089MB
SRB+RRb+ResNet34	88.86%	89.36%	88.24%	88.57%	89.67%	60.014MB

TABLE IV
COMPARISON OF THE TEST RESULTS OF VARIOUS NETWORKS AND THE PROPOSED NCE-NET USING THE PCB-2-DET DATASET

Network	Accuracy	Precision	Recall	NP	Specificity	Param.
SENet	85.83%	87.26%	86.59%	84.16%	84.93%	81.863MB
Inception-v4	87.29%	88.72%	87.69%	85.65%	86.82%	157.201MB
NLNet	86.67%	86.62%	89.27%	86.73%	83.56%	82.184MB
GCNet	87.50%	89.11%	87.74%	85.65%	87.21%	81.529MB
MSDSM	85.21%	86.26%	86.59%	83.94%	83.56%	83.484MB
ECANet	86.25%	87.45%	86.06%	84.98%	86.46%	81.484MB
EPSANet	88.33%	88.10%	**90.80%**	85.39%	88.63%	287.053MB
FCANet	84.17%	85.65%	81.28%	82.88%	86.94%	128.344MB
PCANet	87.50%	87.55%	87.55%	87.45%	87.45%	135.336MB
RMCSAM	86.25%	86.90%	84.68%	85.66%	87.76%	97.696MB
NCENet(Ours)	**88.86%**	**89.36%**	88.24%	**88.57%**	**89.67%**	**60.014MB**

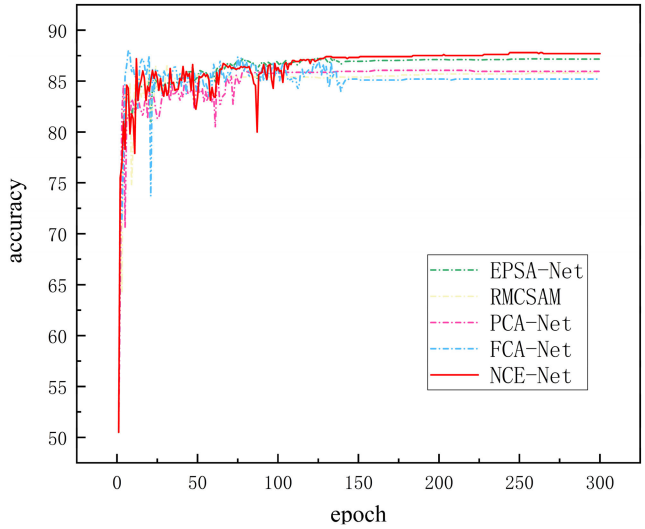

Fig. 8. Accuracy comparison of different models.

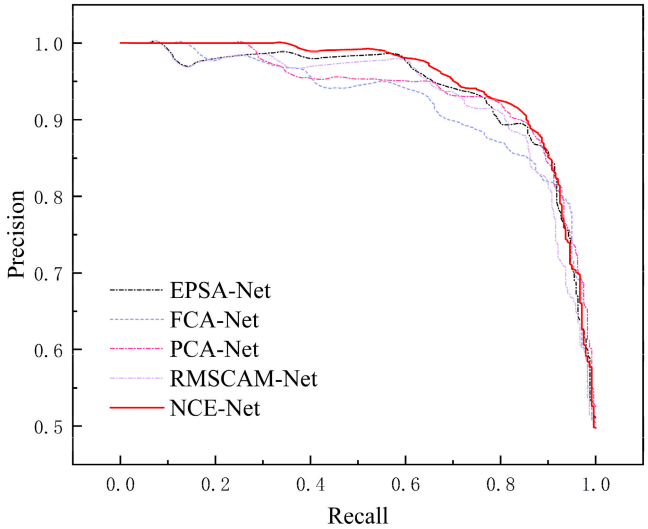

Fig. 9. Comparison of PR curves of different models.

The comparison results also show that most of the current attention mechanisms do not effectively notice the defect features, and our method focuses on guiding the relationship between network attention defects and their neighbors to achieve better classification.

E. Anti-Noise Sensitivity Experiment

In order to prove the robustness of the model proposed in this article in practical detection applications, we conducted anti-noise sensitivity experiments to test the resolution of the model to data with different levels of noise. As shown in Fig. 10, the abscissa represents the variance of Gaussian noise added to the image data, and the ordinate represents the values of Accuracy, Precision, Recall, NP, and Specificity under different levels of noise. It can be seen from the figure that the accuracy of the model continues to decline slightly with the enhancement of the noise level. Obviously, our model is robust to noise to a certain extent.

F. SI Method

Because in industrial applications, enterprises have high requirements for the detection rate of pseudo defects in PCB

TABLE V
SIX DIFFERENT TRAINED NCE-NET VERSIONS

Network	Accuracy	Precision	Recall	NP	Specificity
NCENet-1	88.54%	87.70%	89.84%	89.41%	87.55%
NCENet-2	88.96%	86.64%	93.65%	91.74%	85.11%
NCENet-3	88.13%	88.21%	88.57%	88.03%	87.66%
NCENet-4	88.54%	87.40%	90.61%	89.82%	86.38%
NCENet-5	88.86%	89.36%	88.24%	88.57%	89.67%
NCENet-6	89.17%	87.26%	92.24%	91.40%	85.96%

TABLE VI
SPECIFICITY CALCULATION RESULTS

\	NCENet-1	NCENet-2	NCENet-3	NCENet-4	NCENet-5	NCENet-6
NCENet-1	C=0	C=0.48	C=0.168	**C=1.168**	**C=1.416**	C=0.8
NCENet-2	\	C=0	**C=1.6**	**C=1.264**	**C=1.44**	**C=1.12**
NCENet-3	\	\	C=0	C=0.96	**C=1.15**	C=0.8
NCENet-4	\	\	\	C=0	**C=1.02**	C=0.72
NCENet-5	\	\	\	\	C=0	**C=1.04**
NCENet-6	\	\	\	\	\	C=0

TABLE VII
NETWORK RESULTS AFTER FUSING THREE DIFFERENT NCE-NET VARIETIES,THE FILTERING RATIO REFERS TO THE PROPORTION OF FILTERED DATA IN THE TOTAL DATA

Network	Accuracy	Precision	Recall	NP	Specificity	Filtration rate
NCE Net-2-4-5	92.59%	92.17%	86.53%	93.07%	80.00%	10%

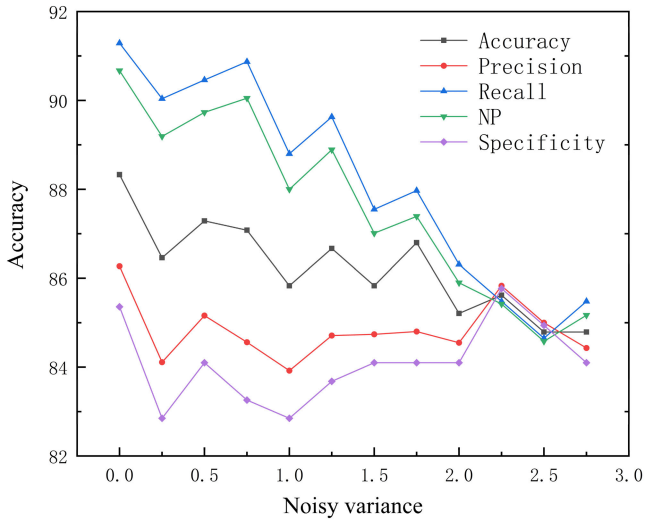

Fig. 10. Robust analysis of image noise.

defect detection, traditional detection methods often add cost-sensitive functions at the model loss calculation end to improve the detection accuracy of pseudo defects at the cost of reducing the detection accuracy of real defects. However, although this method ensures high reliability of pseudo-defect identification, the cost is huge, which often leads to a large number of reusable pseudo defects being judged as real defects. In order to solve this problem, we start from the essence of the PCB true and pseudo-defect detection task. The essential purpose of this task is to reduce the labor cost of the PCB recheck terminal as much as possible while ensuring the highest accuracy of judgment by introducing deep learning. Therefore, we have designed a method that can not only ensure a very high direct defect recognition rate, but also greatly reduce manual work, that is, we have designed a model specificity integration method.

The SI algorithm is a method to calculate the specificity of two models. The greater the specificity, the greater the difference between the two models. On the contrary, the more similar the two models are.

First, we generated six NCE-Net models with peak validation set accuracy after multiple trainings as shown in Table V. Then we use the SI algorithm to calculate the specificity between two models, as shown below

$$C = \frac{m}{k} * \left\{ \frac{\rho(\mathrm{A},\mathrm{B})}{\mathrm{num}(\mathrm{A} \cup \mathrm{B})} * \frac{1}{\rho(\mathrm{A},\mathrm{B})} * \sum_{\mathrm{i}=1}^{\rho(\mathrm{A},\mathrm{B})} |\mathrm{A}_i - \mathrm{B}_i| \right\} \tag{7}$$

$$\rho(\mathrm{A},\mathrm{B}) = \mathrm{num}(\mathrm{A} \cup \mathrm{B}) - \mathrm{num}(\mathrm{A} \cap \mathrm{B}) \tag{8}$$

where A is the sample corresponding to the error judgment of Model 1, B is the sample corresponding to the error judgment of Model 2, A ∪ B is the union of all error judgment samples in Models 1 and 2, A ∩ B is the intersection of all error judgment samples in Models 1 and 2, num () calculates the total number of samples, Ai is the predictive value corresponding to the error judgment sample of Model 1, and Bi is the predictive value corresponding to the error judgment sample of Model 2. m and k are user-defined thresholds, with m representing the standard ratio of A ∩ B to the part where A and B do not intersect. In this experiment, we set it to 2. k represents the standard value of the absolute value of averaged Ai and Bi

and was set to 0.25 for this study. Finally, if the specificity C of the two models is greater than 1, the two models can be regarded as different models, and the two models can be selected for common defect detection. When two models predict that a sample is a pseudo-defect category at the same time, the sample is recognized as a pseudo defect, or when two models predict that a sample is a true defect category at the same time, the sample is recognized as a true defect, otherwise, the sample will be filtered out for manual detection. Table VI shows our calculated specificities, according to which NCENet-2, NCENet-4, and NCENet-5 are mutually different models. These three models were then used as the model group, with their test set calculation featured in Table VII. From this table, precision and NP achieve higher detection at the cost of recall. And only 10% of the samples need to be filtered, which proves the effectiveness of the SI method.

VI. Conclusion

To enhance the classification and detection of true and pseudo defects in PCB recheck terminals, we propose a lightweight high-precision detection method: NCE-Net. In order to maintain the high detection rate of this model, we propose an SRB module and an RRB module. It also includes the LEL module, which is used to improve the sensitivity of the model to the location relationship, and the RC module, which can capture the key relative geographic information in the target map. Our proposed method can effectively enhance the relationship information between defects and their surrounding environment to improve defect classification accuracy. We also put forward a model SI method, which can realize the combination of multiple models to further improve the classification of true- and pseudo-PCB defects. Finally, we present the PCB-2-DET dataset. Experimental results show that our NCE-Net is a suitable target classification method with better accuracy than SOTA when using PCB-2-DET.

In the future, our work can contribute to the design of more effective defect capture modules and improve integration methods, which can further heighten NCE-Net's own defect detection accuracy.

Acknowledgment

The authors thank the Guangdong Greatsense Intelligent Equipment Company Ltd., for providing test data samples and verifying the algorithm studied in this article in the actual production work.

References

[1] J. Jiang, J. Cheng, and D. Tao, "Color biological features-based solder paste defects detection and classification on printed circuit boards," *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 2, no. 9, pp. 1536–1544, Sep. 2012, doi: 10.1109/TCPMT.2012.2205149.

[2] L. Wang, Y. Zhao, Y. Zhou, and J. Hao, "Calculation of flexible printed circuit boards (FPC) global and local defect detection based on computer vision," *Circuit World*, vol. 42, no. 2, pp. 49–54, May 2016, doi: 10.1108/CW-07-2014-0027.

[3] V. H. Gaidhane, Y. V. Hote, and V. Singh, "An efficient similarity measure approach for PCB surface defect detection," *Pattern Anal. Appl.*, vol. 21, no. 1, pp. 277–289, Feb. 2018, doi: 10.1007/s10044-017-0640-9.

[4] X. Zhou, "A surface defect detection framework for glass bottle bottom using visual attention model and wavelet transform," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2189–2201, Apr. 2020, doi: 10.1109/TII.2019.2935153.

[5] P. Wei, C. Liu, M. Liu, Y. Gao, and H. Liu, "CNN-based reference comparison method for classifying bare PCB defects," *J. Eng.*, vol. 2018, no. 16, pp. 1528–1533, Nov. 2018.

[6] B. Ghosh, M. K. Bhuyan, P. Sasmal, Y. Iwahori, and P. Gadde, "Defect classification of printed circuit boards based on transfer learning," in *Proc. IEEE Appl. Signal Process. Conf. (ASPCON)*, Dec. 2018, pp. 245–248, doi: 10.1109/ASPCON.2018.8748670.

[7] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," 2017, *arXiv:1709.01507*.

[8] J. Fu et al., "Dual attention network for scene segmentation," 2018, *arXiv:1809.02983*.

[9] H. Zhang, K. Zu, J. Lu, Y. Zou, and D. Meng, "EPSANet: An efficient pyramid squeeze attention block on convolutional neural network," 2021, *arXiv:2105.14447*.

[10] Y. Miao, Z. Liu, X. Wu, and J. Gao, "Cost-sensitive Siamese network for PCB defect classification," *Comput. Intell. Neurosci.*, vol. 2021, pp. 1–13, Oct. 2021, doi: 10.1155/2021/7550670.

[11] B. Fang, X. Long, F. Sun, H. Liu, S. Zhang, and C. Fang, "Tactile-based fabric defect detection using convolutional neural network with attention mechanism," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–9, 2022, doi: 10.1109/TIM.2022.3165254.

[12] B. Xia, J. Cao, and C. Wang, "SSIM-NET: Real-time PCB defect detection based on SSIM and MobileNet-V3," in *Proc. 2nd World Conf. Mech. Eng. Intell. Manuf. (WCMEIM)*, Nov. 2019, pp. 756–759, doi: 10.1109/WCMEIM48965.2019.00159.

[13] N. Zeng, P. Wu, Z. Wang, H. Li, W. Liu, and X. Liu, "A small-sized object detection oriented multi-scale feature fusion approach with application to defect detection," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–14, 2022, doi: 10.1109/TIM.2022.3153997.

[14] J. Balzategui, L. Eciolaza, and N. Arana-Arexolaleiba, "Defect detection on polycrystalline solar cells using electroluminescence and fully convolutional neural networks," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, Jan. 2020, pp. 949–953, doi: 10.1109/SII46433.2020.9026211.

[15] Z. Ling, A. Zhang, D. Ma, Y. Shi, and H. Wen, "Deep Siamese semantic segmentation network for PCB welding defect detection," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022, doi: 10.1109/TIM.2022.3154814.

[16] Y. Bao et al., "Triplet-graph reasoning network for few-shot metal generic surface defect segmentation," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2021, doi: 10.1109/TIM.2021.3083561.

[17] Q. Lin, J. Zhou, Q. Ma, Y. Ma, L. Kang, and J. Wang, "EMRA-Net: A pixel-wise network fusing local and global features for tiny and low-contrast surface defect detection," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–14, 2022, doi: 10.1109/TIM.2022.3151926.

[18] A. Raj and A. Sajeena, "Defects detection in PCB using image processing for industrial applications," in *Proc. 2nd Int. Conf. Inventive Commun. Comput. Technol. (ICICCT)*, Apr. 2018, pp. 1077–1079, doi: 10.1109/ICICCT.2018.8473285.

[19] Z. Ling, A. Zhang, D. Ma, Y. Shi, and H. Wen, "Deep Siamese semantic segmentation network for PCB welding defect detection," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022, doi: 10.1109/TIM.2022.3154814.

[20] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 3–19, doi: 10.1007/978-3-030-01234-2_1.

[21] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," 2017, *arXiv:1711.07971*.

[22] Q. Hou, D. Zhou, and J. Feng, "Coordinate attention for efficient mobile network design," 2021, *arXiv:2103.02907*.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[24] A. G. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.

[25] J. Dai et al., "Deformable convolutional networks," 2017, *arXiv:1703.06211*.

[26] A. Howard et al., "Searching for MobileNetV3," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 1314–1324, doi: 10.1109/ICCV.2019.00140.

[27] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: More features from cheap operations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1577–1586, doi: 10.1109/CVPR42600.2020.00165.

[28] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the impact of residual connections on learning," *Proc. AAAI Conf. Artif. Intell.*, vol. 31, no. 1, Feb. 2017, pp. 4278–4284, doi: 10.1609/aaai.v31i1.11231.

[29] Y. Cao, J. Xu, S. Lin, F. Wei, and H. Hu, "GCNet: Non-local networks meet squeeze-excitation networks and beyond," 2019, *arXiv:1904.11492*.

[30] R. Ding, C. Zhang, Q. Zhu, and H. Liu, "Unknown defect detection for printed circuit board based on multi-scale deep similarity measure method," *J. Eng.*, vol. 2020, no. 13, pp. 388–393, Jul. 2020, doi: 10.1049/joe.2019.1188.

[31] H. Zhang, L. Jiang, and C. Li, "CS-ResNet: Cost-sensitive residual convolutional neural network for PCB cosmetic defect detection," *Expert Syst. Appl.*, vol. 185, Dec. 2021, Art. no. 115673, doi: 10.1016/j.eswa.2021.115673.

[32] W.-Y. Wu, M.-J.-J. Wang, and C.-M. Liu, "Automated inspection of printed circuit boards through machine vision," *Comput. Ind.*, vol. 28, no. 2, pp. 103–111, May 1996.

[33] W.-C. Wang, S.-L. Chen, L.-B. Chen, and W.-J. Chang, "A machine vision based automatic optical inspection system for measuring drilling quality of printed circuit boards," *IEEE Access*, vol. 5, pp. 10817–10833, 2016.

[34] Z. M. Zhu, Z. K. Lv, and R. C. Song, "PCB solder joint quality inspection based on machine vision," *Control Eng. China*, vol. 27, no. 4, pp. 20–25, 2017.

[35] S. Guan et al., "Ceramic ring defect detection based on improved YOLOv5," in *Proc. 3rd Int. Conf. Comput. Vis., Image Deep Learn. Int. Conf. Comput. Eng. Appl. (CVIDL ICCEA)*, May 2022, pp. 115–118.

[36] G. Liu and H. Wen, "Printed circuit board defect detection based on MobileNet-Yolo-Fast," *J. Electron. Imag.*, vol. 30, no. 4, Jul. 2021, Art. no. 43004.

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

[38] Z. Qin, "FcaNet: Frequency channel attention networks," *Proc. IEEE/CVF Int. Conf. Comput. vision.*, Oct. 2021, pp. 783–792.

[39] D. Liu, Y. Wang, K. Mase, and J. Kato, "Recursive multi-scale channel-spatial attention for fine-grained image classification," *IEICE Trans. Inf. Syst.*, vol. E105.D, no. 3, pp. 713–726, Mar. 2022.

Wenkai Huang (Member, IEEE) received the B.S. and M.S. degrees from Guangdong University of Technology, Guangzhou, China, in 2004 and 2007, respectively, and the Ph.D. degree from Guangzhou University, Guangzhou, in 2017.

In 2007, he joined the School of Mechanical and Electrical Engineering, Guangzhou University, where he is currently an Associate Professor. His research interests include robot vision, medical image processing, and soft robotics.

Jiafu Wen is currently pursuing the bachelor's degree in robotics engineering with Guangzhou University, Guangzhou, China.

His current research interests include deep learning (DL), object detection, attention mechanism, and multi-level feature fusion.

Weiming Gan is currently pursuing the master's degree in machinery engineering with Guangzhou University, Guangzhou, China.

His current research interests include industrial data mining recognition and analysis, deep learning (DL), and image processing technology.

Lingkai Hu received the bachelor's degree from the School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou, China, in 2021, where he is currently pursuing the master's degree in mechanical engineering.

His current research interests include deep learning (DL), object detection, attention mechanism, and multi-level feature fusion.

Bingjun Luo was born in Guangdong, China. He received the B.S. degree from Guangdong University of Technology, Guangdong, in 2004, and the dual M.S. degree from Ecole Polytechnique de Nantes, Nantes, France, and Guangdong University of Technology, Guangdong, in 2006 and 2007, respectively.

He became a Chief Executive Officer of Guangdong Greatsense Intelligent Equipment Company Ltd., Guangdong, in 2012. His research interests include high-end inspection equipment systems.