COMPREHENSIVE REVIEW OF PRINTED CIRCUIT BOARD DEFECT DETECTION USING IMAGE PROCESSING, MACHINE LEARNING, AND DEEP LEARNING TECHNIQUES:

Madeshwar R

B.E - Electronics and Communication engineering

College of Engineering,Guindy,Anna university

ABSTRACT

The fabrication of printed circuit boards (PCBs) is a crucial step in the manufacturing of electronics, and errors in this process can result in broken equipment and large financial losses. Automated flaw detection systems are necessary since traditional manual inspection techniques are labor-intensive and prone to human mistakes. This paper provides a thorough analysis of current developments in PCB defect detection methods, with an emphasis on the fusion of deep learning, machine learning, and image processing approaches.

Keywords: Defect detection, PCB, image processing, machine learning, deep learning

I. INTRODUCTION

One of the most important components in the electronic industry is the printed circuit board (PCB) [1]. It is essential to electronic equipment since it both electrically and mechanically joins different electronic components. As long as integrated circuits and other electronic components are present, printed circuit boards (PCBs) are utilized in practically every type of electronic equipment, including computers, smart phones, watches, and military missile systems. Electronic device components have become extremely small, thanks to advancements in integrated circuit and semiconductor technologies [2]. The PCBs that hold these components are getting smaller, more sensitive, and more sophisticated. To satisfy client standards, they must be produced to a high standard.

It is typically difficult to maintain quality control in the PCB manufacturing process since a range of flaws invariably arise from improper handling or technological errors. Common PCB flaws, including breakout, open circuit, under-etching, mouse bite, spur, short, spurious copper, over-etching, and broken hole, are depicted in Fig. 1.

The increasing complexity and miniaturization of electronic devices have heightened the demand for reliable PCB inspection methods. Traditional visual inspection methods are inadequate for detecting subtle defects in modern PCBs, necessitating the development of automated solutions. This review surveys the state-of-the-art techniques that leverage image processing, machine learning, and deep learning algorithms to enhance defect detection accuracy and efficiency.

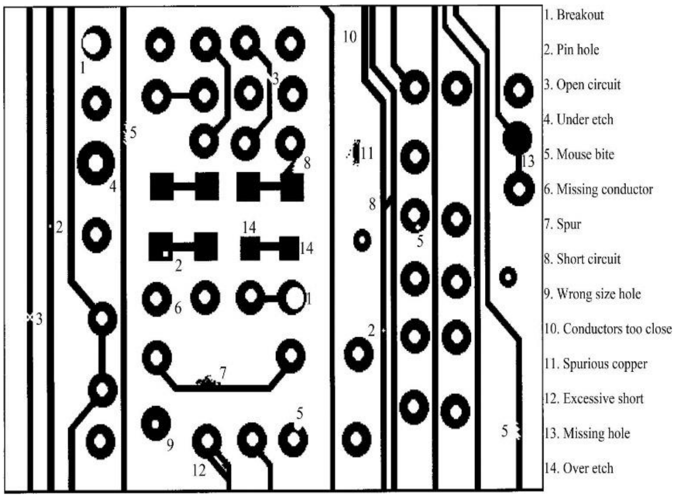

FIGURE 1. Example of defective PCB patterns.

Printed Circuit Board Assembly (PCBA), Soldering, component placement, and screen printing are all part of the production process. Defects could include empty solder (i.e., inadequate solder on the solder point), excess solder (i.e., spilled solder on the solder point), bridge (i.e., the solders are melted together), missing or misaligned components, and faulty solder junctions. Fig. 2 displays examples of PCBA solder junctions that are faulty. Any of these flaws could harm the board as a whole.

Recently, automated optical inspection (AOI) techniques have been routinely employed to detect defects on the PCB assembly line [7]. AOI offers various advantages over traditional manual inspection. For example, it can scan and recognize the board quickly while remaining highly accurate.

This method uses the correlation or difference between the template photos and the inspected images to assist individuals estimate PCB faults. Furthermore, given the non-reference method, a predefined mathematical algorithm is offered for detecting PCB faults. The hybrid technique combines both reference comparison and non-reference procedures.

Previously, fundamental electric tests, such as defect detection, were typically accomplished through manual vision, which needs a small number of trained laborers, necessitating a significant investment in hiring or training them. However, even the most seasoned worker may make mistakes during the inspection. Because components and solder junctions are shrinking and possible flaws are increasing, traditional manual inspection was phased out.

The Fourth Industrial Revolution (Industry 4.0) [3] has brought with it both opportunities and problems for the PCB manufacturing industry. High-quality, precise, and reliable automated industrial processes is the foundation of Industry 4.0 [4]. PCB development must therefore proceed quickly since the manufacturing process for small, complicated PCB boards needs to be more stable, dependable, and quick [5]. Defect detection in PCBs is therefore essential. Many manufactured boards are likely to be discarded later if recurring PCB problems cannot be found quickly and accurately, which is inefficient and expensive. For manufacturers to increase production and profit, real-time, high-precision defect detection and quality control in PCBs are currently essential [6].

Although the AOI system can help factories release a certain number of labourers, PCB industries still need to devote adequate manpower for quality inspection in cooper- ation with the AOI system. Thus, to increase the detection accuracy and speed, which helps decrease labour costs, many researchers focus on building advanced traditional rule-based image processing or machine vision-based detection algo- rithms. In recent years, especially since 2012 when AlexNet was proposed [14], some researchers have tried to apply con- volutional neural networks (CNNs) [15] to extract features for defect detection in PCBs.

Compared to conventional feature extraction techniques, CNN-based models have produced exceptional results in visual detection tasks, including semantic segmentation, object identification, and picture classification. It can accurately record fault features even in the presence of shadows or reflections without the need for extra information. Because of these evident advantages, the CNN-based object detection algorithm consistently updates historical records in most object detection contests and is steadily taking the lead in the industry [16]. Thus, classical image processing-based, machine learning-based, and CNN-based algorithms for PCB and PCBA defect identification will be covered.

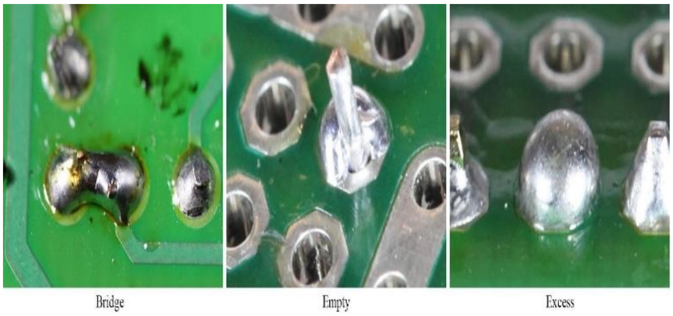

FIGURE 2. Examples of defective solder joints.

This review paper on PCB defect detection makes several significant contributions:

1. It meticulously outlines the systematic methodology employed for collecting data from relevant articles. This includes defining research parameters, selecting appropriate search keywords and databases, and employing classification and verification processes to identify relevant papers. The paper also delves into the research analyzes necessary to synthesize findings from existing PCB defect detection studies.

2. It provides a comprehensive review of PCB defect detection methods over the past two decades, encompassing traditional image-based approaches, machine learning techniques, and deep learning methodologies. Notably, the authors highlight the scarcity of recent reviews in this field, especially those covering deep learning-based methods, as opposed to older reviews.

3. The paper conducts detailed analyses, including the principles and implementation procedures of previous works, to uncover their strengths and limitations. This level of scrutiny surpasses that of previous review papers, which tended to provide more finite assessments.

4. The performance metrics of previous works, such as detection speed and precision, are thoroughly discussed and compared. Importantly, the paper emphasizes the significance of detection speed, alongside precision, in real-world PCB industries, where practicality is crucial. This differs from previous reviews, which often focused solely on detection precision without considering practical deployment factors.

5. Additionally, the paper addresses open research challenges from various perspectives, drawing insights from the detailed analyses conducted on the current state-of-the-art in PCB defect detection.

II. MANUAL DEFECT DETECTION METHODS

Prior to the application of the AOI technique, the majority of manufacturers conducted all electrical testing, including PCB board flaw detection, by hand vision inspection. When determining if a PCB is qualified or failed, and when a correction operation is required, operators always utilize a calibrated microscope or magnifying glass [17]. They use qualified experiences and visual inspection as their tools. Expert inspectors are able to accurately identify problems and their types in this way. The fact that this approach doesn't require test fixtures and has minimal upfront expenses is another advantage.

III. IMAGE PROCESSING-BASED DEFECT DETECTION

In contrast to manual inspection, the AOI system can quickly and quantitatively estimate while eliminating subjective errors. Because it never gets weary or burns out, this system can operate continuously, saving a factory a significant amount of money on labor expenses. Recently, a number of researchers have presented machine vision technology using conventional rule-based image processing method to detect PCB faults in order to enhance detection speed and accuracy, which means better quality and cheaper costs. This section describes algorithms or techniques for PCB or PCBA flaw detection that are based on image processing.

ASPECTS OF SURFACE DEFECTS IN PCB TECHNIQUES

In 2011, Pal et al. [22] used machine vision and the image subtraction approach to identify bare PCB flaws. This method of image subtraction [8], [23], which falls under the category of reference comparison, was carried out in multiple steps. First, one must choose a reference image from a collection of high-quality templates. The reference template needs to be buffered after that in order for subtraction to be permitted. The examined image is then loaded. Pixel-by-pixel XOR logic operation, also known as image subtraction, was used between the reference image and the examined image to determine the PCB faults or errors. If the inspected board had flaws, the image obtained after this subtraction process should show several of them.

MANUAL DEFECT DETECTION METHODS

Prior to the application of the AOI technique, the majority of manufacturers conducted all electrical testing, including PCB board flaw detection, by hand vision inspection.

When determining if a PCB is qualified or failed, and when a correction operation is required, operators always utilize a calibrated microscope or magnifying glass [17]. They use qualified experiences and visual inspection as their tools. Expert inspectors are able to accurately identify problems and their types in this way. The fact that this approach doesn't require test fixtures and has minimal upfront expenses is another advantage. But the effectiveness of this approach mostly rests on the operators' abilities, who should have practice.

IMAGE PROCESSING-BASED DEFECT DETECTION

In contrast to manual inspection, the AOI system can quickly and quantitatively estimate while eliminating subjective errors. Because it never gets weary or burns out, this system can operate continuously, saving a factory a significant amount of money on labor expenses. Recently, a number of researchers have presented machine vision technology using conventional rule-based image processing method to detect PCB faults in order to enhance detection speed and accuracy, which means better quality and cheaper costs. This section describes algorithms or techniques for PCB or PCBA flaw detection that are based on image processing.

Ma [24] proposed a similar method to detect defects in bare PCBs using the comparison between the reference image and inspected image. The modification was the standard image was an average of a series of qualified PCBs instead of using only one image.. Furthermore, a filter based on target region contours was proposed to capture the real defects from these possible regions, which was reported to outperform those of depending on simple image thresholding. The defect types were determined by the changing times of the peripheral bound- ary’s grey value.

TECHNIQUES FOCUSING ON COMPONENT AND SOLDER JOINT DEFECTS

Unlike previous methods on detecting surface defects in PCBs, Sundaraj [23] used the background subtraction method to inspect the missing or misaligned components in PCBA. The author provided a solution that was able to monitor a larger region in a PCB at any moment. This automatic visual inspection system was considered efficient at decreasing the inspection time and increasing the handling capacity. In this work, Mixture-of-Gaussian distribution [31] was chosen for background modelling. Then, three stages—background learning, parameter estimation and pixel classification—were implemented to build a Gaussian density function. During the experiment, the non-defective PCBA was firstly fed to the learning process to acquire the template or background image by capturing about 100 frames of the PCBA board. Subsequently, the image of a defective PCBA with missing or misaligned components was captured. The defective image was subtracted from the background image to obtain a binary mask that presents foreground and background pixels. A large area of foreground pixels indicated some kind of defect on the position.

The method reported in [23] could detect missing or mis- aligned capacitor and chip with an accuracy of more than 90%. However, this method has several limitations. The pro- cess of background learning to build a reference image cost huge time (about 3 to 4 seconds), which makes this tech- nique not suitable for real-time detection. Similar to previous works, image subtraction requires both the reference image and tested image must be placed in the same orientation precisely, which is trivial to be achieved in PCB inspection industry. Besides, the accuracy highly depends on hardware such as a rigid optical table, an advanced camera with a macro zoom lens and a fixed lighting with a suitable illumination angle. Furthermore, this system was captious for inspected images. It failed to obtain good results on the component whose colour was the same as that of the background. Same colour presents same or similar pixel values, which will be considered as background and missed in the subtraction operation.

Conventional normalised cross-correlation (NCC) was used to detect defects by comparing the reference image and defect image in [32], but this method is time consuming as it computes the NCC depending on two-dimensional images. Therefore, Annaby et al. [33] proposed an improved nor- malised cross-correlation (INCC) method to detect integrated circuit (IC) defects such as misaligned components in PCBs. As to the standard image and defect image, considering two corresponding pixels that they contain, blocks with the same size, in which the top left corners are the two pixels, were created. Then, these two-dimensional block images were converted into one-dimensional feature vectors in the match- ing process. Next, these vectors were augmented by spatial statistical features following discrete cosine transform. After this process, the feature descriptors were obtained. As in the classical NCC method, the correlation coefficient between the corresponding one-dimensional feature descriptors in the standard and defect images was then calculated. A pixel with a correlation coefficient smaller than a specific threshold is classified as a defect pixel in the testing image. Otherwise, the pixel is a defect-free pixel.

The INCC algorithm proposed in [33] was compared with other popular methods, including conventional NCC, Yoo2010 [34] and Fouda2015 [35] using PCB image datasets. The results proved that INCC performed better than others. INCC have fewer false alarms than other two methods, and the defect regions are much more coincident than the methods whose defect regions are smaller than those of the standard model. INCC detected defect regions at least four times faster than NCC when the image size was bigger than 400 × 500.

I. MACHINE LEARNING-BASED DEFECT DETECTION

Image subtraction-based traditional image processing methods can achieve high detection accuracy for PCB defects. However, using these methods is always time consuming, and many reference images are essential and can affect the detection results. Many researchers have proposed several machine learning-based algorithms to overcome these difficulties. Many classical machine learning algorithms have been developed and used in PCB defect inspection. In this section, several machine learning algorithm-based methods that mainly focus on solder joints and components are introduced and discussed.

Yun et al. [39] used SVMs and a tiered circular illumination technique to inspect the PCB solder joints. The illumination equipment consists of a Charge-coupled Device (CCD) camera with three circular lamps whose colours were blue, red and green. The solder joint surface has the same nature as the mirror surface. Three lamps illuminate the surface of solder joints at different angles, and then the CCD camera can receive reflected light with different intensity patterns due to the difference in surface slope. The visual information of the solder joints surface can be inferred from the highlight patterns and images. Then, six characteristic features based on average intensity value and percentage of highlights of images from three colour frames were extracted. Finally, each solder region represented by a six-dimensional feature vector was fed into an SVM layer that consisted of four SVM classifiers to output one of several pre-defined types, including excessive, good, insufficient and no solder.

In [39], 402 solder joints were collected to train and test the model. All the insufficient and no solder defects were detected correctly, and for the excessive class and good class, the detection rates were 96.07% and 98%, respectively. This model was compared with other methods, including K-means classifier [40] and backpropagation (BP) classi- fier [41], to show its better performance. However, this paper just achieved solder joints classification without localisation as SVM is a classifier only. The solder joint images with a size of 50 × 100 pixels instead of whole PCB images were fed into the model. This means that in real PCB defect applications, solder joints images must be extracted firstly, which is a tedious and inefficient process. Moreover, how to choose the hyperparameter C and kernel in SVM is a tricky problem. Meanwhile, the SVM is difficult to be implemented for large-scale training samples, because it uses quadratic programming to solve the support vector, in which the cal- culation of M-order matrix will be designed. When the order of matrix is large, it will consume huge machine memory and operation time. In addition, this method required complex three-colour circular illumination system which is not easy to operate.

Ko and Cho [42] combined an Learning Vector Quantiza- tion (LVQ) neural network [43] and fuzzy logic [44] scheme to inspect solder joints. The solder images were also obtained by a three-colour circular illumination system. Firstly, the obtained solder images were divided into three sub-images (right, centre and left columns in the longitudinal direction of a solder joint), which were each fed into three LVQ classi- fiers. The purpose of this operation was to decrease the bur- den and computation, as increased class number decreases the classification performance markedly [45], [46]. The output values of nodes for each classifier were marked from 1 to 10 for right and left columns, and from 1 to 8 for the centre column based on previous studies [47], [48]. Then, the input images were labelled by an expert into five classes based on their experience. Compared with the linguistic approach, the LVQ classifier has difficulty distinctly categorising complex boundaries among various categories. Thus, a fuzzy rule was proposed based on experts' knowledge to obtain better classification results.

The performance of this method in [42] was tested by using 96 images. The total success rate achieved 95.83% among five solder joints types, which was much higher than that of the original LVQ classifier. Similarly, limited by the comput- ing power, this method also got single solder joint images instead of entire PCB images as inputs, which is definitely not suitable for real PCB industry. A complex three-colour circular illumination system was implemented to obtain color

Belbachir et al. [49] proposed a PCB defect inspection system using wavelet transform (WVT) [50] and multi-layer perceptron (MLP) [51] NNs. Complicated illumination lamps were replaced by a common lighting source. The CCD cam- era was moved instead of the PCBs so that the system was more flexible and could be placed on an automatic produc- tion line. A databank that stored all the components and corresponding defect images to detect was built. A reference board was obtained by assembling images of the components from the database based on the circuit layout. A defective board was obtained by overlapping the defect image with the qualified images of components. For the training path, a series of NNs associated with specific components was trained using a set of images from the databank to recognise the corresponding defects that could appear. For the testing path, template matching was applied to identify the reference point, where the CCD was moved to frame the region of interest (RoI). Then, the WVT was implemented on each captured image to extract features, which were fed into a corresponding trained NN. The outputs of the NN displayed the defect class.

With the system in [49], the authors achieved missing component defect classification and faulty solder joint defect classification. Compared with previous works, this method did not require complex illumination system at all. However, a database storing all the components and corresponding defect images must be constructed firstly, which has to take large amount of time and labour. The training dataset was synthesized by components in the database according to the given circuit scheme. The defective circuits (i.e. missing component) could be obtained by overlapping the component with an area having the same color as the background. All of these complicated pre-processing operations make the detec- tion not easy to be deployed in real PCB defect inspection industry. In addition, template matching was implemented in this method leading to the delay of detection speed. More concrete classification metrics and comparison with other methods were not exhibited in the paper.

On the basis of MLP NN and geometric-wavelet (GW) feature extraction, Acciani et al. [52] presented a novel sol- der joint defects inspection method. The method consists of four steps: image acquisition, image pre-processing, fea- ture extraction and classification. The image that contained more than one board was acquired by a CCD camera and then processed by three image segmentations. To extract the target central board, template matching with a real ref- erence image was applied using the horizontal energy and the vertical energy vectors of a greyscale image. Then, the central board from the previous procedure was matched witha real IC image to obtain the IC region in the target image. Similarly, the solder joint images were acquired by template matching using the corresponding reference image. In the feature extraction step, geometric features [53] and wavelet features [54], [55] were used together to form the third feature type named GW features. Finally, these features were sent into the MLP NNs to obtain the classification results.

The method in [52] achieved overall 98.8% classification rate among five solder joint defect types, which was bigger than that of the method that used single geometric feature [56] or wavelet feature. The performance of this method was also progressive compared with that of the LVQ NN with GW features. However, this method used reference PCB images to get target solder joints through several image matching steps, which represents a mass of computations, causing a low detection speed. Obviously, it must have the limitations same as previous traditional image processing-based methods that require reference images. Besides, it is very difficult for MLP to select the number of hidden nodes. The authors applied n-fold cross-validation strategy to decide the hidden neurons without regard to the time cost. Furthermore, only defects classification without localisation was executed in this work. Crispin and Rankov [57] achieved PCB component detec- tion by integrating GA and template matching.

Cai et al. [59] formulated the solder joint inspection issue as an optimisation problem using robust principle component analysis (RPCA) [60], [61]. IC solder joint images were obtained using a three-colour circular illu- mination system similar as that in [39] and [42]. On the basis of RPCA theory, the hue channel of a colour solder joint image, which contained major information for colour perception [62], was decomposed into a low-rank compo- nent and an error component. An appearance model was built by averaging low-rank components of qualified sol- der joints images. Then, according to prior human knowl- edge and appearance, defect score was defined to evaluate the quality of inspected images. A discrimination scheme was applied to make the final decision for a qualified or unqualified image by setting a threshold. The threshold was determined by the maximal value among the defect scores obtained at the iterative stage of building the appearance model.

In the work [59], the best performance of the inspection of unqualified solder joints was achieved when at least 350 qual- ified solder joint images were used to build the appearance model. In that case, the precision, recall and F-measure were 95.65%, 100% and 97.77%, respectively. However, this method only achieved qualified and unqualified classification without localisation. In addition, the building of an appear- ance model was essential in this method, resulting in a very slow speed because of the large computation costs of singular value decomposition for the data matrix at each iteration in the RPCA process. It is sensitive to very small rotations, as this method was based on the hypothesis that solder joint images of the same type have the same size without rotation. Furthermore, a complex three-colour illumination system was required, and prior human knowledge was in incorporated into the defect score to achieve better inspection performance. All of these limitations indicate that this method is not generalized and not suitable for real PCB inspection applications.

Hongwei et al. [63] used improved AdaBoost [64] and DT to classify six types of solder joint defects. The whole inspection process consisted of the training stage and the test stage. In the training stage, the solder joint image was firstly divided into several sub-regions, which had different colour distributions for defective solders. 30 geometric features based on three colours were extracted in each sub- region. Then, an improved AdaBoost was proposed to seek three necessary optimal features according to their classifi- cation performances. Finally, the classifier was built for each sub-region after training. In the detection stage, three steps (sub-region division, critical feature extraction and classifica- tion for regions) were executed based on the training results. For defect diagnosis, a new DT that combines classification and regression tree was implemented to obtain the final defect class.

The method proposed in [63] achieved an overall clas- sification rate of 97.3% without localisation. The average time for one chip was 8.6 milliseconds, which was slower than image comparison method. In this method, only the chip images with a small size instead of whole images were fed into the classification algorithm, and one component contained only one solder joint defect. Thus, this method will be still time consuming when detecting an entire PCB image with various types of components. Similar to the limitation of previous works, a three-colour hemispherical circular illumination was required to infer the characteristics of the solder shape. The improved AdaBoost was used in this work to select fewer but optimal features. But two thresholds for each weak classifier were introduced and a threshold of relevance information was also applied to guarantee the selected classifier brought new information. Extra threshold parameters caused more computation cost, resulting in a slow classification speed.Wu et al. [65] from the same group in [63], achieved solder joint classification using Bayes and SVM. Similarly, a three-colour illumination system was used to obtain solder joint image

Wu and Zhang et al. [69] used colour grads and Boolean rules to inspect solder joint defects. The solder joint images were acquired by the same illumination as [63], [65], and [67]. On the basis of the colour distribution and the solder land size, several regions were defined in the solder joint image. Then, the area features, the barycentre features, the distribution features and the colour grads' features were defined and extracted. Models of 8 different solder joint types, namely, acceptable solder joint, lacked solder, exces- sive solder, pseudo-solder, component shifted, component lift and tombstone, no solder and missed component, were built according to their appearances and previous feature analysis. The relationship between the solder joint types and regions was derived, and Boolean rules were built to classify solder joints.

The method proposed in [69] achieved an average classi- fication rate of 97.7%, which was higher than that of other approaches in [52], [70], [71], and [72]. However, the prereq- uisite is that judging criteria must be constructed according to the solder joints' appearances and various feature analysis. This process deeply depends on the expert experience and prior knowledge, which is time consuming inevitably. Thus, this method is not suitable for large-scale applications in real PCB inspection industry. Furthermore, the requirement for illumination system is critical to the construction of rules as different illumination situation will bring large errors into the forming of criteria. The running time reached 11 seconds for one PCB, which was also much faster than other references in the paper, but it was still far from real-time detection.

Li et al. [73] proposed a semantic segmentation method to inspect components in the PCB by using a random forest pixel classifier. In this work, depth images of PCBs were used as inputs due to its high robustness to the influence of the illumination and texture of the target [74]. Depth images were applied to detect human body pose [75], [76], [77] and hand gesture [78], [79], [80], [81] by pixel classification approaches. A three-dimensional model of the PCB was used to synthesise the training set by computer graphic rendering, which contained the depth images and corresponding colour- labelled images. Pixel samples in the depth images were selected randomly, and their depth difference features were extracted. Pixel classification labels were extracted from the colour-labelled images.

Then, the difference features were used to train a ran- dom decision forest classifier [82], [83], which built a mapping relation between the difference features and the pixel classification labels. Finally, the semantic segmen- tation of PCB components was achieved by the trained random forest pixel classifier. Missing or misplaced com- ponents were also detected by this method. The prediction of the depth images with missing or misplaced components was compared with the colour-labelled images of qualified PCB. Then, the missing or misplaced component can be detected by calculating and analysing the accuracies of each component.For the method in [73], the overall component recognition accuracies were 98.96% and 83.64% on the synthetic and real depth images, respectively. This method performed poorly for real PCB depth images, maybe because the real PCB images have more complicated layout and various noises, which can make the random forest pixel classifier overfitting. Additionally, this method could only classify defective or non-defective caused by missing or misplaced without the exact defect types. Another limitation is that several exper- iments must be performed to choose the optimal training parameters and random seed for the random forest classifier. Hence, this method must be ineffective for real component defects recognition. Machine learning algorithm-based defect detection works are tabulated in Table 4. Most of them focused on solder joints or components defects detection and achieved good performances successfully. However, large gaps between them with real PCB inspection application in the industry still exist. In machine learning-based methods, features sent into the machine learning algorithms are mainly extracted by traditional image processing operations such as edge detection, morphological processing, various image thresholding techniques. Hence, the detection performance of machine learning-based methods deeply depends on the quality of previous extracted features. Meanwhile, each method was designed for special applications, indicating a poor general- ization ability and robustness. The main limitations can be summarised as follows:

- All the proposed methods only achieved classification without localisation. Then main reason is that most machine learning algorithms were developed for classification issues. The localisation of defects can be achieved by adding several image post-processing operations.Image pre-processing were required before the final classification process, and some methods required a qualified template to locate the reference positions or regions of interest. Similar to traditional image processing-based methods, they also have difficulty acquiring reference images, which is time consum- ing and complicated. Their performances were deeply affected by the reference image and prior image processing.
- Another common limitation is that some of them used a complex three-colour illumination system to obtain target images with abundant features, which was hard to implement and thus expensive.
- Due to the influence by many image processing steps and the choices of various parameters in machine lean- ing algorithms, all of these methods cannot perform real-time inspection of a PCB board automatically, although some of them achieved high-speed inspection for one component.Obviously, machine learning-based PCB defects detection methods cannot discard traditional image processing as

To overcome these limitations, deep learning-based algo- rithms is a good choice. Various CNN models have been used in object detection, image segmentation etc. with desirable detection accuracy and speed simultaneously. Features are not extracted by manual designed extractors, but extracted by CNNs automatically. In addition, they are robust to the environment and noise. Thus, PCB defects inspection also can be considered as an object detection issue in nature, which is possible to be overcome by deep learning methods.

II. DEEP LEARNING-BASED DEFECT DETECTION

Recently, deep learning-based methods, especially CNNs, have been widely used in image processing [84], object detec- tion and segmentation [85]. Unlike traditional machine vision methods, CNN-based approaches are able to automatically extract image features, simplifying the image pre-processing process so that the detection accuracy and speed can be promoted effectively. In addition, CNN-based methods are robust to the environment and noise. Even though shadows or reflections exist, significant object detection outputs can still be achieved because of these methods' multi-level fea- ture extraction ability. With these advantages, CNN-based object detection algorithms have outperformed competing algorithms on different datasets and have become the primary driving force for the development of object detection. There- fore, deep learning algorithms have attracted the attention of researchers for PCB defect detection and achieved good detection performance.

The most frequently used and modified CNNs to detect PCB defects are Faster Region-based Convolutional Net- work (Faster R-CNN) [86] and You Only Look Once (YOLO) series. Faster R-CNN is a two-stage object detection model that was developed based on R-CNN [87] and Fast R-CNN [88]. YOLO series is a one-stage object detection network that has both a high detection speed and good accu- racy. In this section, some CNN-based methods proposed by researchers for defect detection in PCBs are presented and discussed.

A. *CNN-BASED METHODS FOR SURFACE DEFECTS*

Ding et al. [89] proposed a tiny defect detection network for PCBs named tiny defect detection network (TDD-net) based on Faster R-CNN. As the PCB images acquired by industrialcamera always have a high resolution and the defect regions always occupy a very small portion compared with the whole image, the traditional Faster R-CNN, which produces anchors using 3 scales and 3 different ratios, is not suitable for tiny defect detection. Thus, inspired by YOLOv2

[90], the author applied k-means clustering to the PCB training dataset bounding boxes to seek reasonable anchor scales automat- ically. Then, data augmentation methods, including adding Gaussian noise, changing light, image rotating, flipping, ran- dom cropping and shifting, were implemented. To maintain the features that have low semantic level, the architecture of Feature Pyramid Network (FPN) [91] was adopted. In FPN, the features with low resolution and strong semantics are connected with features that possess high resolution and weak semantics by lateral connections from top to down, and it pre- dicts features on each level of the pyramid network.

To mitigate the impact on the detection accuracy caused by the small and imbalance dataset, online hard example mining [92] was applied in the whole training phase to improve the quality of RoI proposals.

In the work [89], 693 images with 6 different defects, namely, missing hole, mouse bite, open circuit, short, spur and spurious copper were synthesized. One image contained several defects of the same type. After training and testing, TDD-net achieved a better performance with 98.9% mean average precision (mAP) than other state-of-the-art methods, including Faster R-CNN with different backbones and FPN. This work achieved excellent detection mAP benefited by the fusion of multiscale features using FPN. However, this algorithm is a two-stage frame and has huge parameters, meaning that the detection speed is obviously slower than that of one-stage network. The use of the full connection layers took up a large amount of parameters. RoI pooling was still used in this work, resulting in the loss of transla- tion invariance of subsequent network features, which would affect the final positioning accuracy.

Moreover, the whole retained RoI proposals by the RoI pooling layer will go through the full connection layer and be calculated separately instead of shared calculation between them, leading to huge repetitive computations. All of these shortcomings makes this method not able to achieve real-time detection in real PCB industry.

Adibhatla et al. [93] used the YOLO model to detect defects in PCBs. In order to achieve fast and accurate detec- tion simultaneously, tiny YOLOv2 [94] network modified from YOLOv1 [95], was introduced in this method. The backbone of YOLOv2 is a Darknet-19 network with 19 con- volution layers and 5 MaxPooling layers, which is smaller than traditional Visual Geometry Group (VGG) [96] network with comparable accuracy. The floating-point computation of Darknet-19 is reduced to about 1/5 of VGG-16 to allow for faster arithmetic.

TABLE 4. Summary of machine learning-based methods.

Work	Concept or approach proposed	Detection criterion	Performance	Advantages	Limitations
[39]	SVMs and a tiered circular illumination system	Classification rate	>96%	High classification rate	- Achieved classification without localisation - Complex illumination - Choices of the kernel and hyperparameters in SVM - Focused on small solder joint images only
[42]	Combined LVQ NN and fuzzy logic scheme	Classification rate	95.83%	A promotion than original LVQ	- Achieved classification without localisation - Complex illumination Focused on small solder joint images only - Highly depended on the expert's experience
[49]	Combined WVT and MLP NNs	NA	NA	Common lighting source	- Only classified missing components and faulty solder joints - Required template matching and reference images - A database must be constructed firstly
[52]	MLP NN and GW feature extraction	Classification rate	98.8%	High classification rate	- Required several image matching and inference images - Achieved classification without localisation - Hard to choose the number of hidden nodes in MLP
[57]	Combined GA and template matching	NA	NA	Achieved resistor detection successfully	- Required template matching and reference images - Time consuming, running time of 39.5 s - Selection of parameters
[59]	RPCA	Precision, recall and Fmeasure	95.65%,100% and 97.77%	High classification rate	- Complex illumination - Sensitive to very small rotation - Appearance model must be built, time consuming - Achieved classification without localisation
[63]	Combined improved AdaBoost and DT	Classification rate	97.3%	Higher success rate than original AdaBoost	- Achieved classification without localisation - Time consuming - Complex illumination - Only focused on the chip images with a small size, and one component contained only one solder joint defect
[65]	Combined Bayes and SVM	Classification rate	100%	High classification rate	- Classified one component with one defect, instead of a PCB board - Complex illumination - Required template matching and reference images - Achieved classification without localisation - Time consuming
[67]	Combined BP NN and GA	Classification rate	98.46%	High classification rate	- Classified one component with one defect, instead of a PCB board - Complex illumination - Poor performance for cold solder - Required template matching and reference images - Time consuming - Achieved classification without localisation
[69]	Used colour grads and Boolean rules	Classification rate	97.7%	Classified 8 types of solder joint defects	- Complex illumination - Time consuming, reached 11 s for one PCB - Highly depended on expert experience and prior knowledge
[73]	Random forest pixel classifier for component recognition	Detection accuracies	98.96% or 83.64%	High recognition rate for synthetic images	- Poor performance for real images - Only classified defective components without the exact defect types - Hard to choose the optimal parameters and random seed - Time consuming

[a] Note: NA indicates that information is not available in the references.

TABLE 5. Summary of CNN-based detection models.

Work	Concept or approach proposed	Detection criterion	Performance	Advantages	Limitations
Works focusing on surface defects					
[89]	TDD-net	mAP	98.9%	- High mAP - Detects 6 defects	- One image contained only one type of defects - Not real time yet - Two-stage methods
[93]	Used the YOLO model to detect defects	Detection accuracy	98.82%	High accuracy	- Not sufficient for unbalanced datasets - Only located defective regions without classification - Focused on small local images cropped from whole PCB images
[16]	Improved Faster RCNN and FPN	mAP, Recall	94.2%, 82.5%	One image contained several kinds of defects	- Not real time yet - Cannot obtain a good result for open defects - Low recall
[102]	YOLOv4-MobileNetv3	mAP, F1-score	98.64%, 97.83%	- Real time - High mAP	- Each image contained only one defect - Focused on cropped images with small size from whole PCB images - Designed balanced dataset manually
[106]	Applied YOLOv5 to detect PCB defects	Detection accuracy	>99%	- Real time - High accuracy	- Only located defective regions without classification - Only detected defects from cropped images with small size
[110]	Modified FCOS	mAP	44.3%	Almost real time for high resolution image	- mAP was lower than that of Faster R-CNN and other models - Small volume data
[114]	CS-ResNet	Sensitivity	89%	- Reduce the influence of imbalanced dataset - Greater sensitivity than ResNet	- Difficult to choose a suitable adjustment factor α for different datasets - Performed binary classification only
Works focusing on solder joint or components					
[115]	YOLOv3 and clustering-based active semisupervised classification	Detection precision	90%	Reduce the workload of labelling	- Not real time - Only detected two defect types - The number of selected clusters was a tricky problem
[116]	Hybrid YOLOv2 and Faster R-CNN with retraining process	Detection precision	99.99%	High precision	- Time consuming, 10–15 s inference time for a high resolution image - Need engineers to be involved - More training time
[117]	Deep learning network and image matching to detect solder paste defects	Classification accuracy	96.4%	Detected six types of solder joints with high accuracy	- Template matching was still required - Two-stage method - Not real time
[119]	LD-PCB and CR-PCB to detect component defects	Detection accuracy	95%	High accuracy for component defects	- Template marching or image comparison was required - Three-stage method - Not real time

* Note: NA indicates that information is not available in the references.

The common approaches used by previous works are CNN networks such as Faster R-CNN and the YOLO series. CNN models are able to automatically extract image fea- tures and simplify image pre-processing so that the detection accuracy and speed can be promoted effectively. Further- more, compared with traditional image processing methods and machine learning methods, CNN-based methods are robust to the environment and noise due to their multi-level feature extraction ability. Moreover, in most cases, refer- ence images can be discarded in deep learning models. Deep learning-based defect detection works are tabulated in Table 5. However, these CNN-based methods still need to be improved. The main limitations are summarized as follows:

- Most of them without template matching or image comparison cannot accomplish real-time detection and produce high accuracy at the same time, especially for high resolution PCB images. The main reason isCNNs should contain more layers so that they are deep enough to extract deep features for precise detection. However, a deeper model means many parameters and heavy computations, leading to much more inference time. A balanced point needs to be achieved or some targets should be sacrificed to attain high detection accu- racy or speed. For example, methods of [16] and [89] based on Faster R-CNN achieved good performances in mAP, but they are far away from being able to perform real-time detection, which is important in real indus- tries.

III. PERFORMANCE COMPARISON AMONG EXISTING METHODS

To make the performance comparison among existing CNN-based methods, a new PCB dataset was built based on 10 high resolution PCB images from the public HRIPCB dataset [89]. Similarly, various defects are produced in each image by Photoshop to build a new dataset containing 800 PCB images with an average resolution of 2,777 × 2,138 pixels. Then, the original high resolution images are cropped into 1981 sub-images with the size of 640 × 640 pixels. Compared with [89], the difference is that at least 2 types of defects existing each image of our dataset. There are maximum 6 types of defects in each image, including missing hole, mouse bite, open, short, spur and spurious copper. Data augmentation techniques such as flipping, rotation etc. are applied to this small dataset. Finally, our training set with 11093 images and validation set with 2774 images are built, respectively.

Using our PCB defect dataset, we compare the detection performances of existing state-of-the-art models, including Faster R-CNN, RetinaNet, SSD, YOLOv3 and YOLOv5. The comparison results are reported in Table 6. The experiments are run on a general computer with one NVIDIA GeForce GTX 1660 Super GPU. All the detection results are acquired at the input image resolution of 2240 × 2240 pixels. According to Table 6, Faster R-CNN with backbone of ResNet101, as a representative of the two-stage object detec- tion models, achieves a mAP of 90.6%. RetinaNet with backbone of ResNet101 performs a mAP of 96.2%, which is 5.6% more than Faster R-CNN. However, the detection speed of RetinaNet has no marked increase. SSD-Lite has the least parameters, resulting in the significant sacrifice in precision. YOLOv3-tiny obtains the highest mAP of 99.4% and fastest speed of 19 frames per second, demonstrating the desirable superiority in PCB defects detection. But all of them do not reach the real-time detection speed, and models are still too weight to be applied in real PCB inspection industry.

IV. RESEARCH CHALLENGES

A. *TRENDS IN PCB DEFECT DETECTION*

As stated earlier in this paper, PCB defect detection methods can be mainly divided into four strategies. The first detection strategy is manual visual inspection, which allows experi- enced experts to capture flaws and types of defects. However, with its low efficiency, high cost and poor robustness, it had become increasingly impractical and is thus being eliminated.

RESEARCH CHALLENGES

Traditional image processing-based methods and machine learning-based methods are not fully automatic and real time. Thus, as mentioned in Section VIII-A, deep learning-based PCB defect detection method is an important research focus in the future. Deep learning-based methods completely depend on the computer for automatic detection, which has the obvious advantages of fast detection speed, high recogni- tion accuracy and robust adaptability. Some researchers have applied deep learning methods to obtain acceptable results compared with traditional methods in PCB defect detection, but studies on deep learning methods are still not abundant. The previous review and introduction indicate that several challenges exist for current deep learning methods used for defect detection in PCBs.

1) OPEN PCB DEFECT DATASET

For the supervised deep learning method, few open datasets of PCB defects are available, especially for component defects and solder joint defects. The existing open datasets mainly focus on component detection or cosmetic defect detection with only one type of defect in one image. Having a large amount of images directly contributes to the performance of CNNs. Thus, having numerous defect images is necessary for the training of CNN models. Also, the process of data acquisition and labelling is time consuming.

2) DETECTION WITH BOTH HIGH ACCURACY AND SPEED

For surface defect detection, without subtraction, most deep learning methods classify inputs into two classes only: defec- tive with location identification and non-defective. The spe- cific defect types are not recognised. In other cases, some works are able to detect defect location and type identifi- cation, but the dataset was designed specially, which means that one image contained only one defect or several defects belonging to one type. Even though a few models such as TDD-net have achieved defect location and type identifi- cation, these models should balance the detection mAP or accuracy and speed. Thus, the difficulty lies in achieving high detection mAP and speed simultaneously. In addition, the number of types is limited to about six, which also needs to be increased.

3) TINY COMPONENT DEFECT DETECTION

With the development of micro-electronic technology, the assembled PCB is becoming increasingly smaller. Hundreds of tiny components can be mounted on a small board, which is difficult for human eyes to inspect. Component defects may include missing and misaligned components. Some researchers have focused on the detection of com- ponent types. However, few studies focus on component defect detection only depending on the deep learning method without image subtraction or matching. One of the possible reasons is that collecting PCBA component defects is more difficult than collecting cosmetic defects. An assembled PCB always has a mass of parts such as resistors, capacitors and chips, thus being a more complicated detection environment for CNN models. Image matching is always applied to reduce the inference time.

SOLDER JOINT DEFECT DETECTION

For deep learning-based solder joint defect detection, current methods always rely on outside assistance, such as image matching or registration and manual confirmation or inspec- tion during the detection process, which means that they are not fully automatic. Without assistance, they can only detect the qualified and defective classification with good performances, thus indicating that the detection classes are limited to a small range, which does not coincide with the real industrial manufacturing process.

V. FURTHER RESEARCH DIRECTIONS

A. OPEN PCB DEFECT DATASET

Having a large number of defect images is indispensable for the training of deep learning algorithms. The acknowledged public synthetic PCB cosmetic defect dataset was published by [89] from Peking University. However, its drawback is that one image contains only one type of defects. Further work, such as using Photoshop, can be performed to enrich this dataset so that one image contains at least two types of defects. In addition, some public datasets of solder joint defects and components defects can be downloaded. After they are modified, labels can be marked by software such as LabelMe [124] or LableImg [125]. During the collection process, data augmentation should be executed to obtain high-volume datasets that include different kinds of PCB and PCBA defects.

B. DETECTION WITH BOTH HIGH ACCURACY AND SPEED

Previous work indicated that the YOLO series algorithms can achieve precise and real-time detection, thus making them applicable in real industrial manufacturing. The advanced YOLOv5, YOLOx [126] or YOLOv7 [127] can be used to detect several types of defects in one image, and the per- formance can be verified through a comparison with other state-of-the-art models. The limitations should be analysed and discussed. An improved model focusing on tiny object detection and model compression can be proposed to achieve better PCB defect detection performance.

A. PCB AND PCBA DEFECT DETECTION

Finally, on the basis of previous ideas that have become reality, a CNN model that is able to detect bare PCB cosmetic defects and PCBA solder joint and component defects simul- taneously with acceptable performance can be proposed. This approach will be much more challenging than implementing only one of them, which means that the model must be robust to different environments and backgrounds. Such a model can greatly benefit the PCB-related manufacturing industry.

CONCLUSION

Quality control in the PCB manufacturing process is usu- ally a critical problem. Various defects, including surface, component and solder joint defects, inevitably appear due to mishandling or technical faults. In this paper, different defect detection methods based on conventional image pro- cessing techniques and deep learning models are reviewed and compared. Traditional image processing-based defect detection methods achieve acceptable detection accuracy, but they are time consuming and sensitive to the environment and the inference image. With its multi-level feature extraction and automatic learning abilities, CNN-based defect detection algorithms have overcome such issues and achieved signifi- cant performances in both accuracy and speed. However, deep learning-based methods still has some limitations, such as the collection of datasets, binary classification and small range of defects. On the basis of such shortcomings, some suggestions are proposed and discussed, which are believed to achieve good results in the future.

REFERENCES

[1] Q. Zhang and H. Liu, ‘‘Multi-scale defect detection of printed circuit board based on feature pyramid network,’’ in *Proc. IEEE Int. Conf. Artif. Intell. Comput. Appl. (ICAICA)*, Jun. 2021, pp. 911–914.

[2] T.-M. Bajenescu, ‘‘Miniaturisation of electronic components and the problem of device overheating,’’ *Electrotehnica, Electronica, Automat- ica*, vol. 69, no. 2, pp. 53–58, May 2021.

[3] U. M. Dilberoglu, B. Gharehpapagh, U. Yaman, and M. Dolen, ‘‘The role of additive manufacturing in the era of industry 4.0,’’ *Proc. Manuf.*, vol. 11, pp. 545–554, Jan. 2017.

[4] N. Karnik, U. Bora, K. Bhadri, P. Kadambi, and P. Dhatrak, ‘‘A com- prehensive study on current and future trends towards the characteristics and enablers of industry 4.0,’’ *J. Ind. Inf. Integr.*, vol. 27, May 2022, Art. no. 100294. D. B. Anitha and M. Rao, ‘‘A survey on defect detection in bare PCB and assembled PCB using image processing techniques,’’ in *Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, Mar. 2017, pp. 39–43.

[5] M. Moganti, F. Ercal, C. H. Dagli, and S. Tsunekawa, ‘‘Automatic PCB inspection algorithms: A survey,’’ *Comput. Vis. Image Understand.*, vol. 63, no. 2, pp. 287–313, 1996.

[6] T. Evgeniou and M. Pontil, ‘‘Support vector machines: Theory and appli- cations,’’ in *Advanced Course on Artificial Intelligence*, vol. 2049. Berlin, Germany: Springer, 2001, pp. 249–257.

[7] E. Grossi and M. Buscema, ‘‘Introduction to artificial neural networks,’’ *Eur. J. Gastroenterol. Hepatol.*, vol. 19, pp. 1046–1054, Jan. 2008.

[8] C. R. Reeves, ‘‘Genetic algorithms,’’ in *Encyclopedia of Database Sys- tems*. New York, NY, USA: Springer, 1993.

[9] L. Rokach and O. Maimon, ‘‘Decision trees,’’ in *The Data Mining and Knowledge Discovery Handbook*. New York, NY, USA: Springer, 2005.

[10] M. Seul, L. O’Gorman, and M. J. Sammon, ‘‘Template matching,’’ in *Practical Algorithms for Image Analysis: Description, Examples*

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘ImageNet classification with deep convolutional neural networks,’’ *Commun. ACM*, vol. 60, no. 2, pp. 84–90, Jun. 2012.

[12] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cam- bridge, MA, USA: MIT Press, 2016. [Online]. Available: http://www.deeplearningbook.org

[13] 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.

[14] K. P. Anoop, N. Sarath, and S. V. V. Kuma, ‘‘A review of PCB defect detection using image processing,’’ *Int. J. Eng. Innov. Technol.*, vol. 4, pp. 188–192, May 2015.

[15] N. Aggarwal, M. Deshwal, and P. Samant, ‘‘A survey on automatic printed circuit board defect detection techniques,’’ in *Proc. 2nd Int. Conf. Advance Comput. Innov. Technol. Eng. (ICACITE)*, Apr. 2022, pp. 853–856.

[16] M. F. Ahmad, N. A. M. Isa, W. H. Lim, and K. M. Ang, ‘‘Differential evolution: A recent review based on state-of-the-art works,’’ *Alexandria Eng. J.*, vol. 61, no. 5, pp. 3831–3872, May 2022.

[17] A. Suhasini, D. K. Sonal, B. G. Prathiksha, B. S. Meghashree, and H. D. Phaneendra, ‘‘PCB defect detection using image subtraction algo- rithm,’’ *Int. J. Comput. Sci. Trends Technol.*, vol. 3, p. 8887, May 2015.

[18] A. Mujeeb, W. Dai, M. Erdt, and A. Sourin, ‘‘One class based feature learning approach for defect detection using deep autoencoders,’’ *Adv. Eng. Informat.*, vol. 42, Oct. 2019, Art. no. 100933.

[19] A. Pal, S. Chauhan, and S. Bhardwaj, ‘‘Detection of bare PCB defects by image subtraction method using machine vision,’’ *Proc. World Congr. Eng. (WCE)*, vol. 2, Jul.

[20] K. Sundaraj, ‘‘PCB inspection for missing or misaligned components using background subtraction,’’ *WSEAS Trans. Inf. Sci. Appl.*, vol. 6,
pp. 778–787, Jan. 2009.
[21] J. Ma, ‘‘Defect detection and recognition of bare PCB based on com- puter vision,’’ in *Proc. 36th Chin. Control Conf. (CCC)*, Jul. 2017,
pp. 11023–11028.
[22] R. A. Melnyk and R. B. Tushnytskyy, ‘‘Detection of defects in printed circuit boards by clustering the etalon and defected samples,’’ in *Proc. IEEE 15th Int. Conf. Adv. Trends Radioelectronics, Telecommun. Comput. Eng. (TCSET)*, Feb. 2020, pp. 961–964.
[23] W. Zhu, H. Gu, and W. Su, ‘‘A fast PCB hole detection method based on geometric features,’’ *Meas. Sci. Technol.*, vol. 31, no. 9, Sep. 2020, Art. no. 095402.
[24] M. Fornaciari, A. Prati, and R. Cucchiara, ‘‘A fast and effective ellipse detector for embedded vision applications,’’ *Pattern Recognit.*, vol. 47, no. 11, pp. 3693–3708, Nov. 2011.
[25] T. Atherton and D. J. Kerbyson, ‘‘Size invariant circle detection,’’ *Image Vis. Comput.*, vol. 17, no. 11, pp. 795–803, Sep. 1999.
[26] M. A. Haq, A. K. Jilani, and P. Prabu, ‘‘Algorithmic scheme for concurrent detection and classification of printed circuit board defects,’’ *Comput., Mater. Continua*, vol. 71, no. 1, pp. 355–367, 2022. [Online]. Available: http://www.techscience.com/cmc/v71n1/45360
[27] R. Akmal, N. Z. Othman, A. Bade, M. Rahim, and I. Amin, ‘‘Con- nected component labeling using components neighbors-scan labeling approach,’’ *J. Comput. Sci.*, vol. 6, no. 10, pp. 1099–1107, Oct. 2010.
[28] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis, ‘‘Background modeling and subtraction by codebook construction,’’ in *Proc. Int. Conf. Image Process.*, vol. 5. Oct. 2004, pp. 3061–3064.