Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105
Volume **11**, Number **1**, February **2026**

STD-PCB: Small Object Detection Algorithm For PCB Defects

Jin Han[1,*], Mei-Juan Zheng[1]

[1]School of Computer Science and Engineering,
Shandong University of Science and Technology, Qingdao 266400, China
skd990313@sdust.edu.cn, 1992895665@qq.com

Nai-Ke Shi[2,*]

[2]School of Materials Science and Engineering,
Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore
kenaishi@163.com

[*]Corresponding author: Jin Han, Nai-Ke Shi

Received October 11, 2024, revised February 4, 2025, accepted May 7, 2025.

Abstract. *Printed Circuit Boards (PCB) are fundamental to electronic devices, with their quality profoundly impacting overall product performance. As electronic devices evolve into thinner and more portable forms, the challenge of inspecting PCB for defects has intensified. Given the minute size of these defects, we introduce STD-PCB (Small Target Detection for PCB), an innovative object detection algorithm specifically designed to address this challenge. First, we incorporate a Multi-scale Fusion Attention (MFA) module, enabling the model to concentrate more effectively on detecting even the smallest defects on PCB. Second, the Spatial Pyramid Pooling with Feature Aggregation (SPPFA) module enhances the diversity of feature extraction by leveraging pooling and convolution operations. When these are fused with the original channel feature maps, the detection performance is further improved. Additionally, we introduce a cross-cascading feature fusion channel that mitigates information loss during processing. Finally, to address the imbalance between difficult and easy samples, we employ a novel loss function, Focal EIOU, which assigns appropriate weight to each sample, ensuring that challenging cases receive adequate attention. Through rigorous experimentation, the STD-PCB algorithm achieves an impressive average detection accuracy of 91.2 % for PCB defects, providing a practical and efficient solution for PCB defect detection in the modern electronics industry.*
Keywords: deep learning, attention mechanism, small target detection, PCB defects

1. **Introduction.** Printed Circuit Boards (PCB) occupy a pivotal position in contemporary electronic devices, serving as the backbone and interconnection platform for electronic components. They seamlessly integrate various elements to enable the functionality of these devices[1]. As electronic technology progresses and application fields broaden, the quality and reliability of PCB have grown increasingly crucial. Despite best efforts, defects during production and assembly are unavoidable, and these imperfections can compromise PCB performance, cause malfunctions, or even damage the boards, leading to significant costs[2]. Hence, PCB defect detection has emerged as a vital step in the manufacturing process, aimed at promptly identifying and rectifying potential issues to ensure the highest standards of quality and reliability.

Traditional approaches to PCB defect detection often rely on manual inspection and electrical testing. Manual methods rely heavily on human visual acuity, which can be prone to fatigue and distraction, increasing the risk of human error. Furthermore, for mass-produced PCB, this method proves inefficient and costly[3, 4]. Electrical testing methods, such as bed-of-nails testing and flying probe testing, involve contact-based evaluations that can inadvertently damage the circuit boards during the detection process. Moreover, their applicability is limited, and they are often effective only for specific types and models of products. Given these limitations, there is a pressing need for more efficient and reliable PCB defect detection methods.

Meanwhile, with the progression of integration processes, circuit boards are becoming increasingly dense and complex, housing a greater number of electronic components, wires, and interconnections within a confined physical space. This heightened integration leads to a proliferation of small target defects, posing a significant challenge to defect detection on PCB.

Contribution. In this paper, we introduce STD-PCB, a novel detection algorithm specifically designed to address the intricacies of small target defects on PCB. The key innovations of our approach are as follows:

(1) MFA Module: We devise a novel MFA module to bolster our algorithm's ability to detect minute targets. This attention mechanism strategically directs the model's focus towards small targets, ensuring they receive the necessary attention during detection.

(2) SPPFA Module: Recognizing the potential for feature loss during information extraction, we refine the SPPF module by incorporating a feature extraction channel. This enhanced channel, fused with the original channel post-convolution and pooling, forms the SPPFA module. Furthermore, we establish a cross-layer cascade feature fusion channel to further amplify the information extraction capabilities, mitigating feature loss and enhancing overall performance.

(3) Focal EIOU Loss: To tackle the imbalance between difficult and easy samples in our detection tasks, we adopt the Focal EIOU loss function to calculate the confidence loss. In detection scenarios, easy-to-classify samples often dominate but contribute limitedly to model training. By down-weighting these easy samples through Focal Loss, we effectively reduce the "noise" they introduce, enabling the model to concentrate more intently on the challenging samples. This focused approach leads to improved classification accuracy and a more robust model overall.

In the process of improvement, this paper studies the problem of small target detection, carries out the design of multi-scale attention module and adds the feature fusion path, so that the algorithm can detect small targets better. Meanwhile, in the data enhancement, rotation and noise injection are used to enhance the generalization ability of the model, and the detection accuracy of the algorithm is finally increased by 3.3%.

1.1. **Realted work.** With the advancement of deep learning technology, machine vision-based defect detection methods have gained widespread adoption in the industrial quality inspection realm, particularly for PCB[5, 6, 7, 8]. These deep learning detection methods can be broadly classified into two categories: two-stage and single-stage detection algorithms. Two-stage detection algorithms follow a sequential approach. Initially, a Region Proposal Network (RPN) identifies a set of potential target regions. Subsequently, these proposed regions undergo refinement within a target detection network for recognition and classification, exemplified by algorithms like R-CNN[9], Fast R-CNN[10], and Faster R-CNN[11]. Two-stage algorithms excel in detection accuracy due to their distinct steps for region proposal and classification, enabling more precise target localization and recognition. However, their computational complexity often results in slower processing speeds,

potentially limiting their real-time applicability. Conversely, single-stage detection algorithms streamline the process by treating target detection as a direct regression problem. This approach integrates target region proposal, classification, and localization into a single pass, exemplified by algorithms like SSD[12] and the YOLO series[13, 14, 15, 16]. Single-stage methods are generally faster, catering to the demands of real-time detection.

Recent advancements have further optimized these algorithms. Chen et al.[5] proposed the Transformer-YOLO algorithm, which incorporates an improved clustering algorithm to generate anchor boxes tailored to the dataset. The algorithm utilizes the Swin Transformer as the feature extraction network and integrates an SE(Squeeze-and-Excitation)[17] attention module, effectively enhancing the model's detection accuracy. Tang et al.[6] introduced a PCB-YOLO algorithm, an adaptation of YOLOv5, to address the issue of low detection accuracy for PCB surface defects. The approach incorporates a Swin Transformer in the backbone network and adds a small target detection layer to enhance the detection performance for small-scale defects. This modification results in significant improvements in defect detection accuracy, demonstrating effective results. Yuan et al.[18] proposed the YOLO-HorNet-MCBAM CARAFE (YOLO-HMC) algorithm, based on YOLOv5. They utilized HorNet as the backbone network for feature extraction and designed the Multiple Convolutional Block Attention Module (MCBAM) to enhance small target detection performance. Additionally, they replaced the upsampling layer with CARAFE, significantly improving detection accuracy. Ding et al. [19] introduced TDD-net, an innovative small target detection network that fuses Faster R-CNN with a feature pyramid network, yielding impressive detection results.

Wu et al.[20] combined a lightweight network with a dual-attention mechanism within the GSC YOLOv5 algorithm, effectively tackling the challenge of small target detection. Wang et al. [21] proposed the AE-YOLO algorithm for PCB detection, utilizing CoT as the backbone network for feature extraction. To enhance the model's ability to capture global information, the SPPFS module was added at the end of the backbone network. Additionally, the CC3 module was integrated to better perceive high-level semantic information. This approach not only improved detection accuracy but also increased detection speed. Li et al.[22] improved the YOLOv3 algorithm by using both real and virtual PCB images as datasets for joint training. They also expanded the output headers, leading to improved detection results.

1.2. **Organization.** Section 2 provides a detailed description of the dataset preparation. In Section 3, we introduce the proposed STD-PCB algorithm and explain the specific improvements made. Section 4 presents the evaluation metrics and analyzes the experimental results. Finally, Section 5 concludes with a summary of the findings.

2. PCB dataset construction.

2.1. **Original PCB dataset.** The dataset utilized in this paper originates from the Intelligent Robotics Open Laboratory of Peking University, comprising a comprehensive collection of six distinct PCB defect types: missing hole, mouse bite, open circuit, short, spur, and spurious-copper. This dataset initially encompasses 693 images[23], as detailed in Table 1, providing a valuable foundation for our research.

2.2. **Data Enhancement.** However, given the relatively small size of the original PCB dataset, concerns such as overfitting and compromised detection accuracy could potentially arise during the training phase. To address these issues, we employed a suite of data augmentation techniques, including rotation, brightness adjustment, and the addition of salt-and-pepper noise. These methods systematically transform and expand the samples

Table 1. Original PCB inspection data set

Defect type	Number of images	Number of defects
Missing hole	115	497
Mouse bite	115	492
Open circuit	116	482
Short	116	491
Spur	115	488
Spurious-copper	116	503
Total	693	2953

within the original dataset, thereby generating a more extensive and diverse dataset. The result of this data enhancement process is a significantly enhanced dataset that now contains 2,079 images with a total of 8,859 defects, effectively mitigating the limitations of the original dataset. Furthermore, to ensure the quality and accuracy of our annotations, we leveraged the Make Sense image annotation tool, which facilitated the generation of precise one-to-one XML files for each image in the expanded dataset. The detailed characteristics of this enhanced dataset are presented in Table 2.

Table 2. PCB inspection dataset after data enhancement

Defect type	Number of images	Number of defects
Missing hole	345	1491
Mouse bite	345	1476
Open circuit	348	1446
Short	348	1473
Spur	345	1464
Spurious-copper	348	1509
Total	2079	8859

2.3. **Division of dataset.** The dataset is typically partitioned into three distinct subsets: the training set, the validation set, and the test set, following a random distribution. In this paper, to ensure the model's effectiveness and generalization capabilities, we adopt a division ratio of 7:1.5:1.5 for the training set, validation set, and test set, respectively. This strategic allocation ensures a balanced representation of data across all subsets while prioritizing the training phase. Figure 1 illustrates the distribution of defects within these subsets post-division, providing a clear visualization of the data composition. Furthermore, Figure 2 showcases an illustrative example of the diverse defect types present within the dataset. This visual representation underscores the complexity and variety of defects encountered in PCB inspection, highlighting the importance of a robust and accurate detection model.

3. **STD-PCB.** Yolov5, an important version of the Yolo series algorithms, has high detection accuracy while taking into account real-time requirements. This makes it suitable for defect detection in industrial scenarios with high real-time requirements. At the same time, Yolov5 supports various export formats such as ONNX and Tensorflow, which makes it easy to deploy. Based on this, this paper chooses Yolov5 as the benchmark model for improvement.

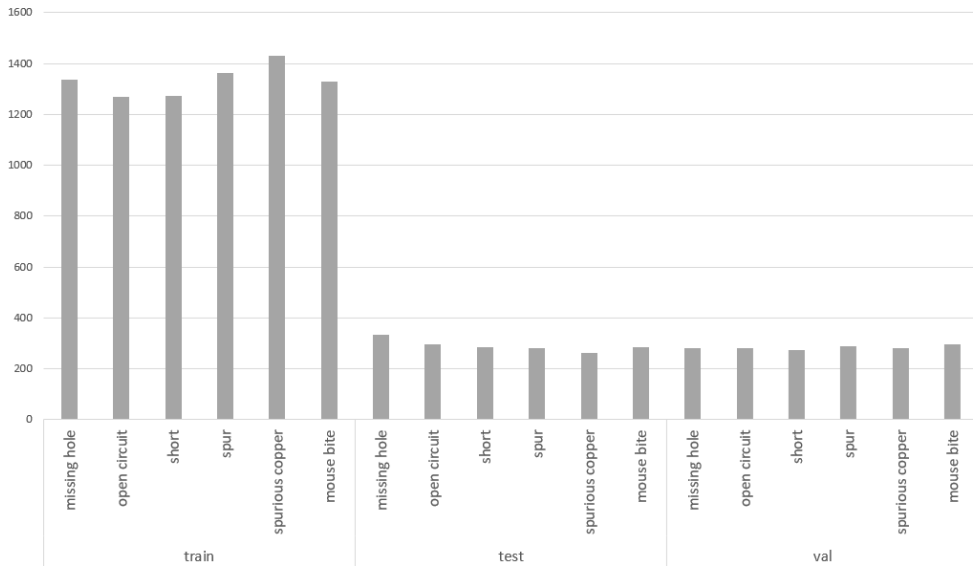

Figure 1. Distribution of defects in different datasets

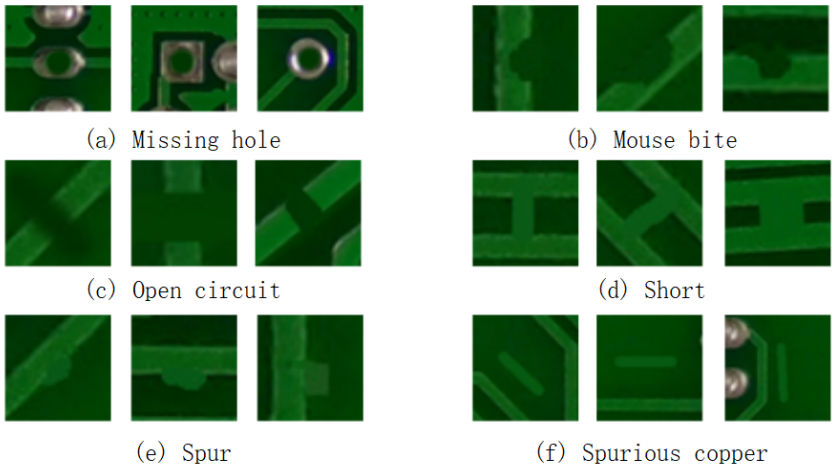

Figure 2. Various defects on the surface of a printed circuit board

3.1. **STD-PCB overall architecture.** This paper introduces a novel small target detection algorithm tailored specifically for PCB defect detection, named STD-PCB. At its core, STD-PCB employs CSPDarknet-53 as the backbone network, leveraging its robust feature extraction capabilities. To enhance the model's sensitivity towards detecting minute defects on PCB, we innovate by introducing the MFA module. This module meticulously focuses the model's attention on the nuanced features of minor defects, significantly improving detection precision. Furthermore, we propose an advanced SPPFA module and integrate CSP[24]-PAN as the neck of the model. This strategic combination facilitates a seamless fusion of deep and shallow semantic features, enhancing the model's comprehension of complex PCB defect patterns. To minimize information loss, we incorporate a cross-hierarchical feature fusion channel between the backbone and neck networks, ensuring optimal utilization of low-level semantic features. To tackle the challenge of imbalanced difficult-to-easy samples during detection, we adopt three distinct scale detection heads as our output layer and employ the Focal Efficient Intersection over Union (Focal EIOU) loss function. This approach effectively addresses the issue,

further refining the model's performance. Collectively, these enhancements culminate in the STD-PCB model, a comprehensive framework designed to excel in detecting even the most minor PCB defects. Figure 3 provides a comprehensive overview of the STD-PCB model's intricate architecture, showcasing its innovative components and seamless integration.

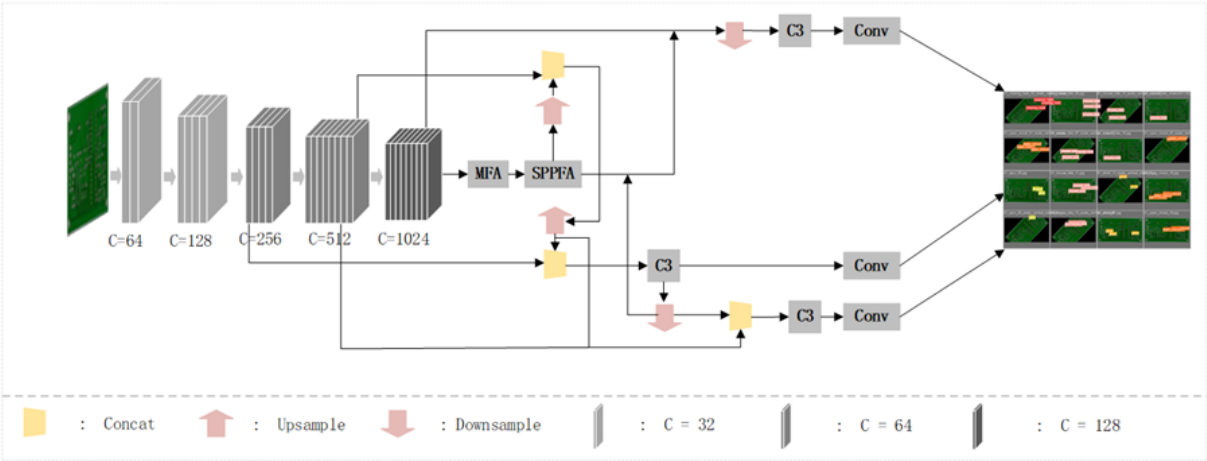

Figure 3. Network structure of STD-PCB

3.2. **MFA Attention Module.** The attention mechanism, a widely adopted weighting approach in deep learning, plays a crucial role in enhancing model performance by effectively directing focus toward task-relevant elements. Its effectiveness is derived from three key aspects: firstly, it enables models to efficiently extract and utilize salient features, thereby amplifying the response in areas of interest. Secondly, attention mechanisms enhance the expressiveness of feature maps by filtering out redundant information and mitigating the disruptive effects of noise. Lastly, by reducing interference from specific channels, attention mechanisms improve the robustness of the model. To address the challenges associated with small target detection in PCB defect analysis, this paper introduces the Multi-scale Fusion Attention (MFA) mechanism. MFA innovatively employs dilated convolutions at multiple scales to extract comprehensive feature information. As depicted in Figure 4, MFA employs a sophisticated weighting strategy, utilizing multiple parameters to meticulously adjust the importance of each channel's features. This nuanced control enables the model to concentrate its attention on the most salient features, crucial for pinpointing minute PCB defects. By implementing MFA, we empower our model to detect even the smallest PCB defects with unparalleled accuracy and reliability. This refinement in detection capabilities underscores the significance of the attention mechanism in enhancing the performance of deep learning models tailored for complex and nuanced tasks such as PCB defect detection.

The MFA module meticulously extracts features from the input feature maps by employing dilated convolutions with dilation rates of 1 and 2, respectively. This dual-scale approach yields feature maps Y1 and Y2, both of identical dimensions. These feature maps then undergo an element-wise summation and subsequent Global Average Pooling (GAP) to consolidate into a single feature map Y. To further enrich Y's feature representation, it is subjected to a 1-dimensional convolutional operation with a kernel size of k. This operation delves deeper into the features, and the resulting map is passed through a sigmoid activation function to normalize its values. Importantly, this 1-dimensional convolutional and sigmoid activation process is repeated on the GAP feature map to ensure a comprehensive feature extraction. Subsequently, the activated feature maps are

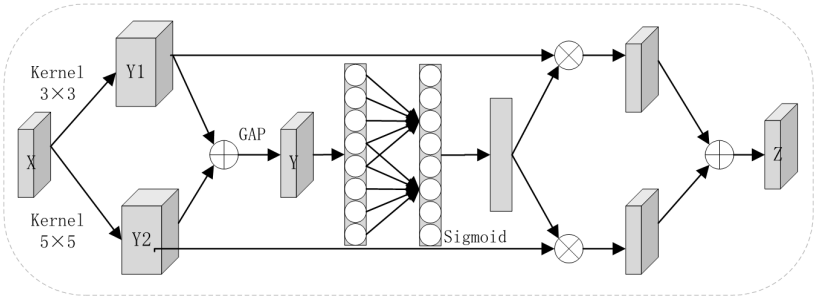

Figure 4. Structure of MFA attention Mechanism

meticulously multiplied, channel-wise, with Y1 and Y2, respectively. This step allows for a fine-grained modulation of the initial feature maps based on the extracted contextual features. Finally, the resulting modulated feature maps are summed element-by-element to yield the MFA module's final output.

The incorporation of dilation in the convolutional kernels significantly enhances the model's ability to comprehend contextual information within the image. By strategically introducing "voids" or gaps in the kernel, dilated convolution expands the receptive field without inflating the parameter count. This not only aids in capturing the global structure and interdependencies of objects within the image but also contributes to reducing the model's computational complexity and improving its efficiency. Figure 5 vividly illustrates the distinction between ordinary and dilated convolution, highlighting the latter's advantages in terms of receptive field augmentation and parameter efficiency.

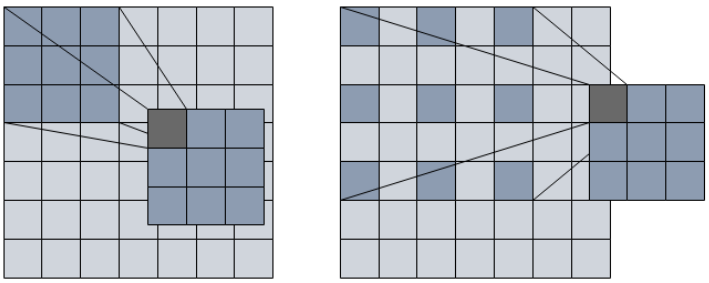

Figure 5. Ordinary Convolution vs. Dilated Convolution

3.3. **SPPFA module with Cross-tier cascade.** The SPPF (Spatial Pyramid Pooling Fast) architecture in YOLOv5, also recognized as the space pyramid pool, is renowned for its ability to adaptively generate outputs of varying sizes. While SPPF efficiently performs both maximum pooling and convolution operations on input feature maps, a potential drawback lies in the maximum pooling step, which retains solely the maximum value within each pooling window. This approach, unfortunately, disregards other valuable data points that might contain crucial detection-relevant information, ultimately leading to a loss of feature richness and potentially impacting detection accuracy. To mitigate this issue, this paper introduces an innovative SPPFA module. This enhancement adds a dedicated feature extraction channel that seamlessly integrates with the original processing channel. Specifically, the input image traverses two parallel paths: one employs the SE (Squeeze-and-Excitation) attention mechanism to meticulously extract salient features, while the other maintains the conventional SPPF processing. The semantic information garnered from these two channels is then meticulously fused and

subjected to convolutional processing, culminating in a refined and informative final result. By incorporating this additional feature extraction channel, the SPPFA module effectively mitigates information loss, ensuring that no potentially crucial detection cues go unnoticed. The resulting improved SPPFA network architecture, depicted in Figure 6, represents a significant advancement in balancing feature richness and detection performance.

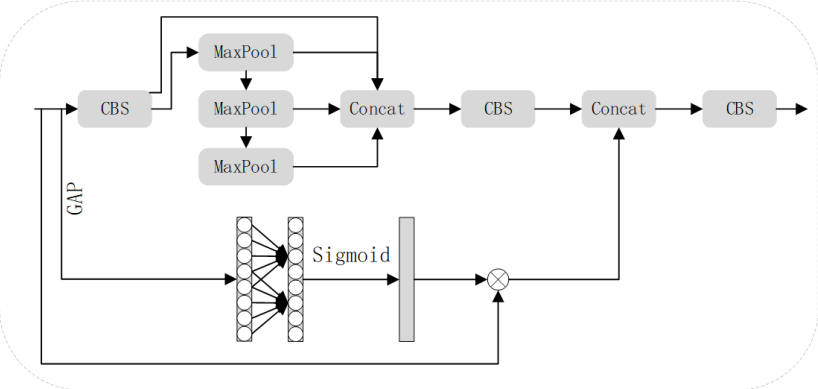

Figure 6. SPPFA structure

This paper employs a Cross-Level Cascade approach to fuse low-level and high-level semantic features to minimize information loss further and enhance the detection of small targets. This method reduces information loss during feature extraction and leverages contextual information from high-level features to capture small target features better. Figure 7 illustrates the structure of the cross-cascade, with the red lines indicating the additional cascade channels. By establishing links across different hierarchical levels, this approach enables the mapping of underlying features to multiple branches, thereby reducing computational and memory requirements. Furthermore, cross-hierarchy linking facilitates the model's ability to learn features at various scales, enhancing its robustness and generalization capabilities. Integrating low-level and high-level semantic features improves the model's ability to detect small targets and increases detection accuracy.

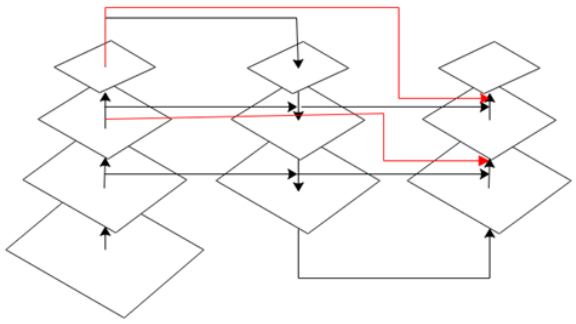

Figure 7. Cross-Level Cascade Schematic

3.4. **Loss function.** The loss function employed in this context comprises three components: classification loss, confidence loss, and bounding box loss. The classification loss, predominantly calculated using cross-entropy, quantifies the closeness between the actual and desired outputs. A lower cross-entropy value signifies a tighter alignment between these two, as depicted in Equation (1). To address the pervasive issue of positive-negative

sample imbalance in datasets, a balancing parameter α is introduced into the cross-entropy loss formula, as seen in Equation (2). While α effectively mitigates the imbalance between positive and negative samples, another challenge arises in the form of a difficult-easy sample imbalance. In this scenario, easy samples contribute minimal loss but have a limited impact on enhancing the model's performance. Conversely, the model should prioritize focusing on difficult samples. To further refine the loss allocation and reduce the impact of high-confidence, easy samples, Focal Loss is introduced. This approach ensures that the model pays greater attention to challenging samples, thereby facilitating a more efficient learning process and enhancing overall model effectiveness.

$$CE = \begin{cases} -\log(p), & if \quad y = 1 \\ -\log(1-p), & if \quad y = 0 \end{cases} \tag{1}$$

$$BCE = \begin{cases} -\alpha\log(p), & if \quad y = 1 \\ -\alpha\log(1-p), & if \quad y = 0 \end{cases} \tag{2}$$

The refined Focal Loss, as outlined in Equation (3), adeptly addresses two critical imbalances: the positive-negative sample imbalance and the difficult-easy sample imbalance. By incorporating the parameter α, it effectively manages the former, ensuring a balanced treatment of samples from different classes. Concurrently, the parameter γ plays a crucial role in addressing the latter challenge by differentially weighting the losses based on the difficulty of defect identification. Specifically, Focal Loss assigns lower weights to defects that are easily identifiable, thereby reducing their impact on the overall loss. Conversely, it assigns higher weights to more challenging defects, thereby encouraging the model to focus more intently on these difficult-to-identify cases. In this way, Focal Loss dynamically adjusts its emphasis, ensuring that the model's learning efforts are optimized to enhance the accuracy of defect identification across the entire spectrum of difficulty.

$$FL = \begin{cases} -\alpha(1-p)^{\gamma}\log(p), & if \quad y = 1 \\ -(1-\alpha)p^{\gamma}\log(1-p), & if \quad y = 0 \end{cases} \tag{3}$$

In the context of Focal Loss, (1-p) falls within the range [0,1], where p represents the predicted probability of the correct class. Here, 1-p can be interpreted as a measure of the distance between the true and predicted values, with a larger distance indicating greater difficulty in learning. Consequently, the loss weight for such difficult-to-recognize samples, modulated by γ, is increased. As γ grows larger, the loss associated with challenging defects becomes more significant, emphasizing the importance of accurately identifying these defects.

When analyzing the detection results of PCB board surface defects, it becomes evident that the model's accuracy for defects like mouse bites and spurs is less than ideal, pointing to a challenging sample imbalance issue. To address this, we opted for Focal Loss as the loss function, leveraging its ability to dynamically adjust weights based on sample difficulty.

In addition, considering the small size of the defects in the PCB defect detection conducted in this paper, CIOU's treatment effect on small targets and sample imbalance is limited, especially when the target position and size difference is large, CIOU's fine adjustment of the border sometimes causes excessive punishment, leading to the instability of the optimization process. EIOU is an extension of CIOU, which is more robust in the detection of small targets and irregularly shaped targets, and can optimize the target frame more accurately. Especially in the face of complex backgrounds and small targets, the effect of EIOU is better than CIOU. The calculation process of EIOU is shown in Equation (4).

$$\begin{aligned} L_{EIOU} &= L_{IoU} + L_{dis} + L_{asp} \\ &= 1 - IoU + \frac{\rho^2(b,b^{gt})}{c^2} + \frac{\rho^2(w,w^{gt})}{c_w^2} + \frac{\rho^2(h,h^{gt})}{c_h^2} \end{aligned} \tag{4}$$

In the context of EIOU, Cw, and Ch represent the width and height, respectively, of the smallest rectangle that encloses the two predicted and ground truth bounding boxes. Notably, EIOU computes the loss in three distinct components, offering a comprehensive assessment of the detection performance. By directly incorporating the edge length as a penalty term, EIOU addresses a common issue where certain edge lengths might be erroneously inflated, ensuring a more accurate measurement of the bounding box's fit.

To tackle both the challenges of difficult-easy sample imbalance and small target localization accuracy, this paper innovatively combines Focal Loss with EIOU to form a hybrid loss function, Focal EIOU. This approach leverages the strengths of both methods: Focal Loss can reduce the interference of background samples to the training process by adjusting the weight of the loss function, so that the model can focus more on learning difficult-to-classify targets, especially small targets and sparse targets. EIOU itself has a strong ability to fit the target frame, especially in small target positioning, which can provide more accurate loss optimization. Combined with Focal Loss, the model can not only locate small targets better, but also avoid being disturbed by a large number of background samples during training. By combining the two, Focal EIOU loss function takes into account the problem of target frame accuracy and sample imbalance in the optimization process, and realizes the more accurate and stable optimization of small targets and difficult to detect targets. The resulting Focal EIOU formula, as presented in Equation (5), provides a powerful tool for enhancing the overall performance of object detection models, particularly in scenarios involving complex samples and fine-grained localization requirements.

$$L_{FocalEIOU} = IoU^{\gamma} L_{EIOU} \tag{5}$$

4. **Experiment and Analysis.**

4.1. **Experimental environment.**

In this experiment, we employed Python as the primary programming language and leveraged the PyTorch framework to facilitate the training process of our model. To optimize the model's parameters efficiently, we chose Stochastic Gradient Descent (SGD) as our optimizer. The experiment was conducted using a Tesla-PCIE GPU to harness its parallel processing capabilities and accelerate the training process. Throughout 100 epochs, the model was trained with a batch size of 16, ensuring a balance between memory utilization and training speed.

4.2. **Evaluation Metrics.**

In this paper, we have carefully selected three key metrics to evaluate the performance of the algorithms: Mean Average Precision (mAP), Precision (P), and Recall (R). These metrics provide a comprehensive understanding of the detection capabilities. Additionally, Intersection over Union (IOU), as defined in Equation (6), serves as a fundamental measure to quantify the overlap between the predicted and ground truth bounding boxes. Precision, outlined in Equation (7), represents the proportion of correctly predicted positive samples among all samples labeled as positive by the algorithm, offering a measure of the algorithm's accuracy in identifying positive instances. Recall, as defined in Equation (8), indicates the fraction of actual positive samples that are correctly identified by the algorithm, assessing its ability to identify all relevant instances within the dataset. To gain a holistic view of detection performance across different categories, we calculate the Average Detection Precision, which integrates both Precision and Recall. This metric, shown in Equation (10), provides an average score

that considers the detection accuracy for each category, offering a balanced assessment of overall performance.

$$IOU = \frac{TP}{FP + TP + FN} \tag{6}$$

$$P = \frac{TP}{TP + FP} \tag{7}$$

$$R = \frac{TP}{TP + FN} \tag{8}$$

$$AP = \frac{\sum_{i=1}^{n} P_i}{n} \tag{9}$$

$$mAP = \frac{\sum_{i=1}^{n} AP_i}{k} \tag{10}$$

In the above formula, TP is the number of positive samples correctly classified, FP is the number of positive samples incorrectly classified, and FN is the number of negative samples incorrectly classified. P can evaluate the accuracy of the model in predicting positive samples, and the higher the value, the higher the accuracy of the model in predicting positive samples.

4.3. **Ablation Experiment.** To evaluate the efficacy of our proposed MFA attention mechanism, enhanced SPPFA module, cross-cascade approach, and loss function, we conducted a rigorous ablation study by comparing the performance of various modules under identical conditions. For this study, we selected the YOLOv5s version as our benchmark algorithm. We employed a methodical approach where we progressively integrated the improved modules to isolate the impact of each enhancement on the detection performance. Throughout the experiment, we maintained a consistent input image size of 640 $\times$ 640 and trained the models for 100 epochs.

The results of our ablation study, presented in Table 3, reveal insightful findings. mAP, Recall (R), and Precision (P) serve as our primary metrics for assessing detection performance. By analyzing the performance gains in Table 3, it becomes evident that both the MFA attention mechanism and the refined SPPFA module contribute positively to enhancing the detection capabilities of the model. Specifically, (i) the integration of the MFA attention mechanism leverages dilated convolutions at multiple scales to enhance feature extraction for small targets. This expansion of the receptive field enables the model to better comprehend the contextual information surrounding small objects, resulting in a notable improvement in mAP from 87.9% to 89.9%. This augmentation underscores the MFA's effectiveness in bolstering the model's detection proficiency. (ii) The refined SPPFA module, which incorporates additional processing channels post-convolution and average pooling, followed by fusion with the feature map after max pooling, significantly enhances performance. This approach increases the mAP from 87.7% to 88.9%. Moreover, to reduce information loss, we introduced cross-layer linkage channels to facilitate seamless information fusion between the trunk and neck components. This enhancement, denoted as P, further elevates mAP to 90.8%, underscoring its role in preserving critical data during the detection process. (iii) To address the challenge of imbalanced samples, particularly in detecting intricate defects such as rat bites and burrs, we implemented the Focal EIOU loss function. This function effectively manages hard-to-detect samples,

improving the model's capability in identifying small targets. As a result, the introduction of Focal EIOU enhances mAP from 87.7% to 88.5%. When these enhancements are combined, the final model achieves an impressive 3.3% increase in mAP compared to the baseline, demonstrating the cumulative effectiveness of our proposed improvements. In the table, we also compared the size of the model with the number of parameters. According to the data in the table, with the improvement of the detection accuracy of the algorithm, the model has become more complex, but it is still within the acceptable range.

Table 3. Comparison of ablation experiments

	MFA	SPPFA	P	Focal EIOU	$mAP/\%$	$P/\%$	$R/\%$	$GFLOPs$	$Parameter$
Base	-	-	-	-	87.9	90	84.3	15.8	7026307
1	✓	-	-	-	89.9	92.8	86	16.0	7290032
2	-	✓	-	-	88.9	92.1	84.8	16.2	7551110
3	-	-	✓	-	90.8	91.7	87.6	16.4	7370371
4	-	-	-	✓	88.5	92.1	83	15.8	7026307
5	✓	✓	✓	✓	91.2	93.4	87.9	17.0	8044681

As illustrated in Figure 8, the STD-PCB algorithm proposed in this paper has been effectively employed to detect defects in PCB. The resulting detection labels and precise imagery demonstrate the algorithm's exceptional performance in defect detection. The test results indicate that the STD-PCB algorithm achieves high accuracy and positioning precision, enabling precise identification and localization of defects in their respective positions.

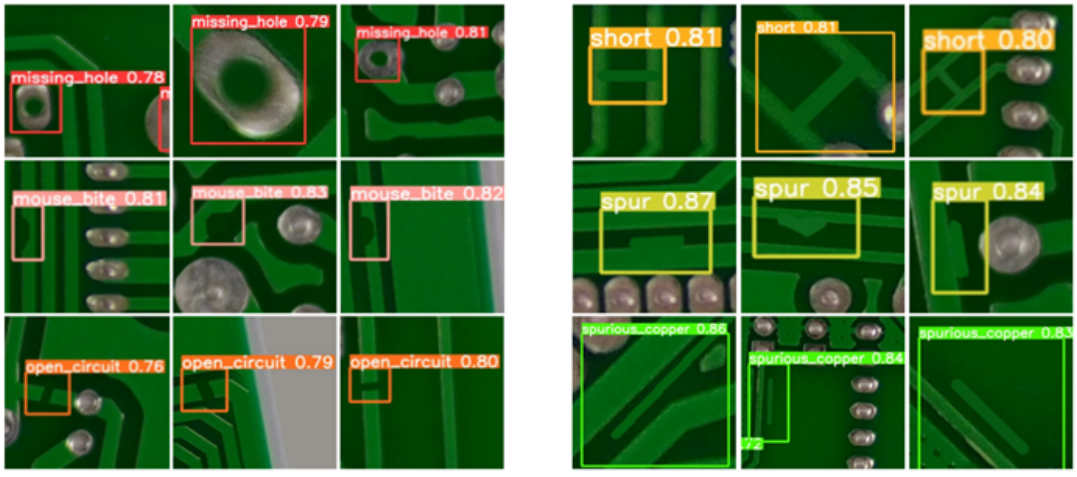

Figure 8. Example of defect detection by STD-PCB algorithm

A statistical analysis of the PCB defect detection results, presented in Table 4, underscores the algorithm's efficacy. Specifically, it achieves an average detection accuracy exceeding 90% for missing holes, open circuits, short circuits, and residual copper, demonstrating its high accuracy and reliability in detecting these four critical defect types. However, we acknowledge that the detection accuracy for rat bite and burr defects lags, likely due to the similarity in their shapes, which can lead to confusion during the detection process. To enhance the detection accuracy further, we plan to delve deeper into effectively distinguishing these two defect types. Potential avenues include augmenting the training

dataset and adopting more sophisticated feThis paper employs a Cross-Level Cascade approach to fuse low-level and high-level semantic features to minimize information loss further and enhance the detection of small targets. This method reduces information loss during feature extraction and leverages contextual information from high-level features to capture small target features better. Figure 7 illustrates the structure of the cross-cascade, with the red lines indicating the additional cascade channels. By establishing links across different hierarchical levels, this approach enables the mapping of underlying features to multiple branches, thereby reducing computational and memory requirements. Furthermore, cross-hierarchy linking facilitates the model's ability to learn features at various scales, enhancing its robustness and generalization capabilities. Integrating low-level and high-level semantic features improves the model's ability to detect small targets and increases detection accuracy.ature extraction techniques to bolster the algorithm's ability to differentiate between these defects. Our ongoing research endeavors will focus on these strategies to refine and optimize the STD-PCB algorithm for even more accurate and reliable PCB defect detection. Meanwhile, the confusion matrix in Figure 9 shows the performance of the model.

Table 4. Detection results of STD-PCB algorithm

	Missing hole	Mouse bite	Open circuit	Short	Spur	Spurious-copper
$P/\%$	98.6	90.3	92.6	94.6	91.5	92.7
$R/\%$	98.9	78.3	85.2	94.4	79.5	91.1
$AP/\%$	98.9	83.7	90.6	96.2	82.4	95.3

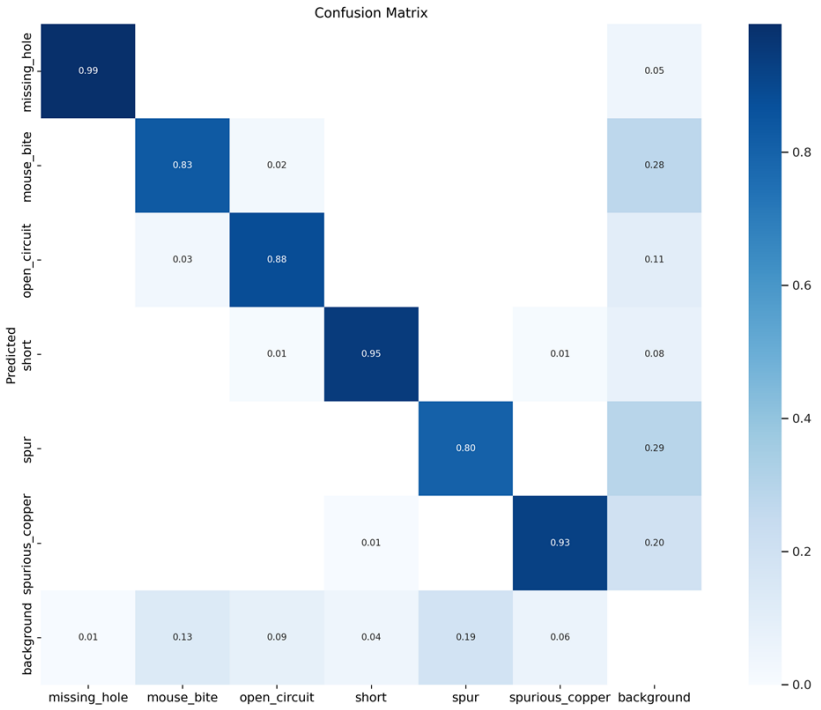

Figure 9. Confusion matrix

As can be seen from the data in Figure 9, although this model performs well in most categories, there are some false positives and missed detection problems in spur, spurious-copper and other defects detection. In the subsequent optimization, we consider to further

improve the model performance by adjusting the threshold and increasing the training of negative samples.

4.4. **Comparison Test.** The attention mechanism, a prevalent weighting approach in deep learning, excels at directing focus toward task-relevant aspects, thereby elevating model performance. In a controlled experimental setup, we compared the performance of our proposed STD-PCB algorithm against both two-stage (Faster R-CNN) and single-stage (SSD, YOLOv7, YOLOv8, and YOLOv3-tiny) target detection algorithms for PCB defect detection. As evident from Table 5, the STD-PCB algorithm significantly outperforms Faster R-CNN, a two-stage approach, in terms of average detection accuracy. Moreover, it surpasses single-stage algorithms like SSD and YOLOv3-tiny by a substantial margin and even maintains an edge over YOLOv5, which leverages multiple backbone networks. Notably, the STD-PCB algorithm also demonstrates improved average detection accuracy compared to the state-of-the-art YOLOv7 and YOLOv8 algorithms within the YOLO family. In summary, the STD-PCB algorithm excels in PCB defect detection tasks, underscoring its ability to enhance the accuracy and reliability of this critical process. This outcome not only validates the effectiveness of the STD-PCB algorithm but also presents a promising solution for the field, offering a new benchmark for PCB defect detection.

Table 5. Comparison experiments of different algorithms

			$AP/\%$					
	Backbone	$mAP/\%$	Missing hole	Mouse bite	Open circuit	Short	Spur	Spurious copper
Faster RCNN	Resnet50	40.97	44.21	36.22	37.68	39.97	44.86	42.89
SSD	VGG16	75.82	89.03	75.3	71.86	84.51	65.28	68.94
YOLOv3-tiny	Darknet-19	67.3	90.3	39.7	60.6	79.2	57.9	76.1
YOLOv5	Resnet18	86.2	97	75.9	81.7	90	81	91.5
	Resnet34	87.5	95.9	78.9	81.9	94.3	80.9	93.2
	MobileNetv2	75.9	98	64	60	87.2	64.9	81.6
	Darknet-53	87.9	99	78.2	84.9	94.4	92.6	78.5
YOLOv7-tiny	-	50.4	89.4	31.6	43.3	62.6	40.5	45.3
YOLOv7[25]	-	89.5	97.3	83.2	89.1	92.8	81.4	93.2
YOLOv8	ShuffleNetv2	91.1	98.5	88.3	90	95.7	83.6	90.2
STD-PCB	Darknet-53	91.2	98.9	83.7	90.6	96.2	95.3	82.4

5. **Summary.** This paper introduces the STD-PCB algorithm, specifically designed to tackle the challenge of detecting small target defects in PCB. The approach begins with the integration of the MFA multi-scale attention module, which enhances the model's ability to detect defects by extracting features across different scales. Additionally, the refined SPPFA module, incorporating cross-layer linkage channels, effectively mitigates information loss and bolsters feature expressiveness. Furthermore, the adoption of the Focal EIOU loss function addresses the issue of imbalanced difficult samples, resulting in a marked improvement in mAP. Experimental results demonstrate that the proposed STD-PCB model significantly outperforms both the single-stage detection algorithm Faster R-CNN and the two-stage detection algorithms YOLOv7 and YOLOv8, achieving an average detection accuracy of 91.2%. This advancement offers a valuable contribution to PCB defect detection methodologies. Faced with the problem of increasing PCB complexity in the future, we will also consider further optimizing the deep learning model to enhance its processing power through more efficient network architecture and optimization algorithms. At the same time, self-supervised learning and small sample learning can be considered, so

that STD-PCB can continue to effectively learn and improve performance without large amounts of labeled data.

Acknowledgment. This work is supported by the Natural Science Foundation of Shandong Province (ZR2021MD057), the Shandong Province graduate education quality course construction project(SDYKC21063), and the Fundamental Research Funds for the Central Universities, China (FRF-TP-22-137A1).

REFERENCES

[1] J. A. Magera and G. J. Dunn, "Printed circuit board," Dec. 2 2008, uS Patent 7,459,202.
[2] Q. Ling and N. A. M. Isa, "Printed circuit board defect detection methods based on image processing, machine learning and deep learning: A survey," *IEEE Access*, vol. 11, pp. 15 921–15 944, 2023.
[3] Y. Li and S. Li, "Defect detection of bare printed circuit boards based on gradient direction information entropy and uniform local binary patterns," *Circuit World*, vol. 43, no. 4, pp. 145–151, 2017.
[4] Y. Deguchi, S. Dobashi, N. Fukuda, K. Shinoda, and M. Morita, "Real-time pcb monitoring using time-of-flight mass spectrometry with picosecond laser ionization," *Environmental Science & Technology*, vol. 37, no. 20, pp. 4737–4742, 2003.
[5] W. Chen, Z. Huang, Q. Mu, and Y. Sun, "Pcb defect detection method based on transformer-yolo," *IEEE Access*, vol. 10, pp. 129 480–129 489, 2022.
[6] J. Tang, S. Liu, D. Zhao, L. Tang, W. Zou, and B. Zheng, "Pcb-yolo: An improved detection algorithm of pcb surface defects based on yolov5," *Sustainability*, vol. 15, no. 7, p. 5963, 2023.
[7] K. Xia, Z. Lv, K. Liu, Z. Lu, C. Zhou, H. Zhu, and X. Chen, "Global contextual attention augmented yolo with convmixer prediction heads for pcb surface defect detection," *Scientific Reports*, vol. 13, no. 1, p. 9805, 2023.
[8] T. Jian, Y. Yang, H. Baoshuai, and H. Chongqing, "Pcb defect detection algorithm based on yt-yolo," in *2023 35th Chinese Control and Decision Conference (CCDC)*. IEEE, 2023, pp. 976–981.
[9] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 580–587.
[10] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1440–1448.
[11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2016.
[12] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, 2016, pp. 21–37.
[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.
[14] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7263–7271.
[15] ——, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
[16] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.
[17] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.
[18] M. Yuan, Y. Zhou, X. Ren, H. Zhi, J. Zhang, and H. Chen, "Yolo-hmc: An improved method for pcb surface defect detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–11, 2024.
[19] R. Ding, L. Dai, G. Li, and H. Liu, "Tdd-net: a tiny defect detection network for printed circuit boards," *CAAI Transactions on Intelligence Technology*, vol. 4, no. 2, pp. 110–116, 2019.
[20] L. Wu, L. Zhang, and Q. Zhou, "Printed circuit board quality detection method integrating lightweight network and dual attention mechanism," *IEEE Access*, vol. 10, pp. 87 617–87 629, 2022.
[21] Y. Wang, Y. Li, D. M. S. Kayes, H. S. Abdullahi, S. Gao, H. Zhang, Z. Song, and P. Lv, "Research on a lightweight pcb detection algorithm based on ae-yolo," *IEEE Access*, 2024.

[22] J. Li, J. Gu, Z. Huang, and J. Wen, "Application research of improved yolo v3 algorithm in pcb electronic component detection," *Applied Sciences*, vol. 9, no. 18, p. 3750, 2019.

[23] W. Huang and P. Wei, "A pcb dataset for defects detection and classification," *arXiv preprint arXiv:1901.08204*, 2019.

[24] C.-Y. Wang, H.-Y. M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "Cspnet: A new backbone that can enhance learning capability of cnn," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 390–391.

[25] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 7464–7475.