Research Article

Research on Online Defect Detection Method of Solar Cell Component Based on Lightweight Convolutional Neural Network

Huaiguang Liu [ID],[1] **Wancheng Ding** [ID],[1] **Qianwen Huang,**[1] **and Li Fang**[2]

[1]*Key Laboratory of Metallurgical Equipment and Control Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China*
[2]*STRE International Engineering Technology Co., Ltd, Beijing 100000, China*

Correspondence should be addressed to Wancheng Ding; dingwancheng_wust@163.com

The defects of solar cell component (SCC) will affect the service life and power generation efficiency. In this paper, the defect images of SCC were taken by the photoluminescence (PL) method and processed by an advanced lightweight convolutional neural network (CNN). Firstly, in order to solve the high pixel SCC image detection, each silicon wafer image was segmented based on local difference extremum of edge projection (LDEEP). Secondly, in order to detect the defects with small size or weak edges in the silicon wafer, an improved lightweight CNN model with deep backbone feature extraction network structure was proposed, as the enhancing feature fusion layer and the three-scale feature prediction layer; the model provided more feature detail. The final experimental results showed that the improved model achieves a good balance between the detection accuracy and detection speed, with the mean average precision (mAP) reaching 87.55%, which was 6.78% higher than the original algorithm. Moreover, the detection speed reached 40 frames per second (fps), which meets requirements of precision and real-time detection. The detection method can better complete the defect detection task of SCC, which lays the foundation for automatic detection of SCC defects.

1. Introduction

Among all kinds of renewable energy, solar energy, as a kind of primary energy of renewable resources, is expected to become the fastest growing renewable energy with its obvious advantages such as clean, safe, and inexhaustible [1]. Solar cell component (SCC) is the key part of photovoltaic power generation system which converts solar energy into electric energy. The quality of the SCC directly affects the output power and service life of photoelectric conversion [2]. In the SCC production process, due to the influence of materials, processes, and human factors, the silicon wafers in SCC will inevitably generate various defects, such as cracks, scratches, and black spots. These defects not only reduce the yield of the SCC but also affect the service life and the efficiency of photoelectric conversion. Therefore, the defect detection of silicon wafer is critical for quality improvement of finished SCC.

At present, there are two defect detection methods of silicon wafers: physical inspection methods and visual inspection methods. The physical detection method relies on personal experience and involves more limitations for efficiency and precision, while the visual inspection method has advantages in accuracy, effectiveness, and stability and is widely used in quality detection [3]. For vision-based solar cell quality detection methods, the imaging schemes mainly include electroluminescence imaging (EL) and photoluminescence imaging (PL). EL technology needs to contact the solar cell for power-on detection, which may cause secondary damage to the cell by electric current, and also has a constrained detection efficiency [4], while the PL combines light sources with different wavelengths to irradiate, so the silicon ion transitions produce luminescence imaging without touching the solar cell. The PL method not only can image the surface and internal defects of the solar cell at the same time but also can detect the process

sheet (nonfinished cell), which is more conducive to the quality control of the product. Therefore, PL technology has gradually become the main imaging technology for solar cell detection [5]. Figure 1 shows the processing of PL imaging.

For solar cell defect detection, Chen et al. [6] proposed a cell crack defect detection scheme based on structure perception. By designing the structure similarity measure (SSM) function, using the nonmaximum value suppression method to extract candidate crack defects, the proposed SSM function has stronger crack defect protrusion and suppression of randomly distributed grains. It is for further extraction of crack defects that provide efficient preparation. Experimental results showed that this method has a good effect on the detection of prominent cracks, but the parameters in SSM cannot be automatically selected, which results in limitations for other types of defect detection. Tsai et al. [7] proposed an automatic defect detection scheme based on haar-like feature extraction and a new clustering technology. In the training process, only images without defects are selected as training samples, and a simple distance threshold is automatically determined for each cluster. Experimental results showed that this method can effectively detect various defects in solar cells, but the algorithm has limitations on the size of the image to be detected. Chen et al. [8] designed a visual defect detection method based on multispectral deep convolutional neural network (CNN). By adjusting the depth and width of the model, the impact of model depth and kernel size on the recognition results was evaluated. Experimental results showed that the multispectral deep CNN model can effectively detect surface defects of solar cells, has higher accuracy and stronger adaptability to large-area defects, but has weak feature extraction capabilities for small-area defects and linear defects. Liu et al. [9] proposed a support vector machine algorithm with radial basis as the kernel function. This method used integral projection and gray barycentric algorithm to obtain the geometric characteristic parameters of solar cell defects and used these parameters as a support vector machine. After learning, the accuracy of the support vector machine to identify common defects was over 90%, but the real-time performance of the detection method needs to be improved. Wang et al. [10] proposed a solar cell surface defect detection algorithm based on deep learning (DL), which reconstructed the image through a built-up deep belief network, and compared the detection result with the real defect image to realize the defect detection of the image. Lichun et al. [11] proposed a solar cell surface quality detection method based on machine vision and artificial neural network in response to the low efficiency and accuracy of solar cell surface quality detection. For the tiny defect of broken grid, the correct recognition rate can reach 98.57% by training the regularized RBF classifier. Finally, the classifier was used in the defect detection system; however, the system cannot identify the internal defects such as hidden cracks and black spots.

The defects are small and blur in the solar cells, and general detection methods will cause false detection and missed detection. In the field of target detection, algorithms based on CNN have the characteristics of strong feature extraction ability and high accuracy. In recent years, they
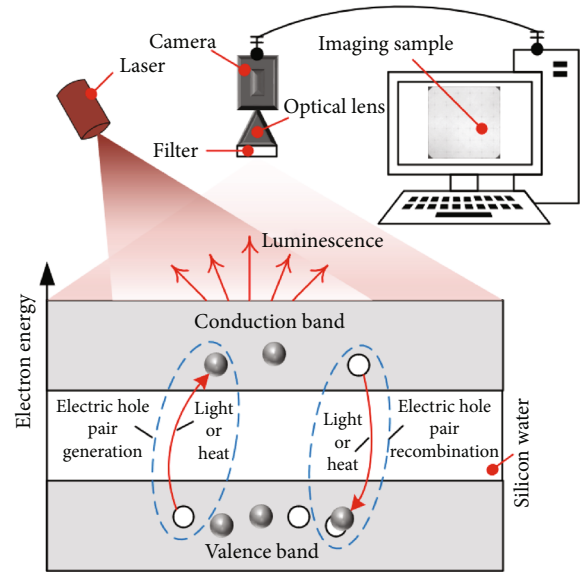

FIGURE 1: The imaging principle of PL technology.

have emerged and been widely used [12]. On the whole, these algorithms can mainly be divided into two branches: two-stage algorithms and one-stage algorithms. The two-stage algorithms generate a series of candidate regions (Region Proposal) through the feature extraction network firstly [13] and then classify and regress the candidate regions on this basis. The R-CNN series of algorithms are typical representation of two-stage algorithms, such as R-CNN, Fast R-CNN, and Faster R-CNN [14–16]. The R-CNN series have higher detection accuracy, but the detection speed is slower due to more network levels. The main idea of the one-stage algorithm is to directly classify and regress the target through the CNN and remove the candidate region step of the two-stage algorithm. Representative algorithms are YOLO [17], SSD [18], and so on. Compared with the two-stage algorithm, the one-stage algorithm has higher real-time performance with relatively lower accuracy. However, with the continuous development in recent years, the one-stage algorithm has a greater improvement in detection accuracy while taking into account the real-time performance, making it easier to meet the industrial application requirements. Therefore, the one-stage algorithm has gradually become the focus of current research. Li et al. [19] used a fully convolutional YOLO detection network to provide an end-to-end solution for strip steel surface defect detection. The network was used to evaluate six types of defects, and a 97.55% mapping rate and 95.86% recall rate were achieved. Qiu et al. [20] developed an autonomous visual detection system based on various defects of wind turbine blades, combined with CNN and YOLO models. In order to detect small-size defects, a small target detection method based on YOLO using multiscale feature pyramid fusion multilayer features is proposed. Experimental results showed that this method was superior to existing methods in detection accuracy and reliability with an average accuracy rate of 91.3%. Compared with other one-stage detection algorithms, the YOLO series has the characteristics
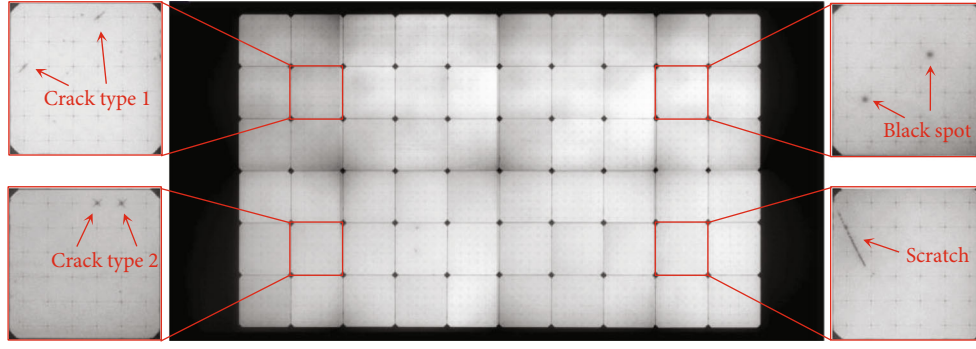
FIGURE 2: PL image of SCC and partial enlarged view of defect.

of fast speed, simplicity and efficiency, high accuracy, and strong generalization ability [21]. As a simplified version of YOLOv3, YOLOv3-Tiny has a simple model structure, lower hardware requirements, and faster detection speed, meanwhile, introduces decreased detection accuracy and increased missed detection rate as the simplified backbone network [22]. Yi et al. [23] aimed at the problem that YOLOv3-Tiny real-time pedestrian detection often loses part of the detection accuracy. Deepen the YOLOv3-Tiny network, enhance the feature extraction ability of the target, and use k-means clustering on the training set to find the best prior box. Experimental results show that this method has high detection accuracy under the premise of meeting real-time performance.

In this paper, defect detection is carried out on the PL image of the SCC. The defect types include the crack type1, the crack type2, scratch, and black spot (as shown in Figure 2). In order to detect tiny defects in high-pixel SCC image, a local difference extremum of edge projection (LDEEP) method is proposed to segment and extract solar cell units in the SCC. As the small size of the defects and the background interference, common one-stage algorithm cannot reach production demands. In view of this, an improved lightweight convolutional neural network model was established. This model deepened the backbone feature extraction network structure with enhancing feature fusion layer, and third-scale prediction layer, thereby strengthening the model performance of feature extraction and small defect target detection. The experimental results proved that the lightweight convolutional neural network model proposed in this paper improves the accuracy of solar cell defect detection under the requirement of real-time performance. This method is suitable for the detection of the crack type1, the crack type2, scratch, and black spot of the solar cells and can complete the detection tasks of the above-mentioned defect types in the SCC. Figure 2 shows PL image of SCC and partial enlarged view of defect.

The rest of this article is as follows: Section 2 introduces the SCC unit segmentation. Section 3 introduces the defect recognition based on lightweight deep neural network. Section 4 introduces the experimental results and analysis. Finally, Section 5 provides the conclusion of this article.

2. Solar Cell Component Unit Segmentation

As the SCC images have more than 70 million pixels (about 7900×4100), whereas defects usually account for small areas (about 64×64) in a single silicon wafer (about 600×600) of the SCC image (as shown in Figure 2). If the SCC image is detected directly, it will cause a large amount of calculation and time-consuming. Meanwhile, the defects occupy relatively small proportion, which is prone to miss detection for the common detection methods. A viable solution under the premise of ensuring efficiency is to extract each single silicon wafer and then transfer defect position to the whole coordinate of SCC image. Under this strategy, it is necessary to perform single silicon wafer segmentation firstly.

2.1. Local Extremum Neighborhood Difference Based on Edge Projection. From Figure 2, it can be seen that the SCC image (with 10×6 silicons) has a black background. In order to obtain each silicon wafer region, the component region should be extracted firstly and segmented with a priori knowledge of layout (10×6 or 12×6). So minimum bounding rectangle (MBR) is adopted for the component region extraction as shown in Figure 3.

As shown in Figure 3, a whole gray SCC image is changed to binary image firstly to divide target and background, and then, the binary image is processed with the MBR to obtain the target region of the SCC silicon wafer. Hence, the position coordinates of the four corners A, B, C, and D of the rectangle are also determined. Assuming that the position coordinates of the first corner point A on the upper left are (x, y), then the coordinates of points B, C, and D are shown in Table 1.

In Table 1, w_s and h_s are the width and height of the image of the silicon wafer region, respectively. Then, segment and extract the SCC image area $[y : y + h_s, x : x + w_s]$ to obtain the silicon wafer.

As the silicon wafer is arranged on the moving board by a robot arm, there are prominently distributed horizontal and vertical lines. In order to extract the straight lines in the image to facilitate subsequent edge projection positioning, the *Sobel* edge detection is used to perform differential operations to highlight the changes of the boundary. The
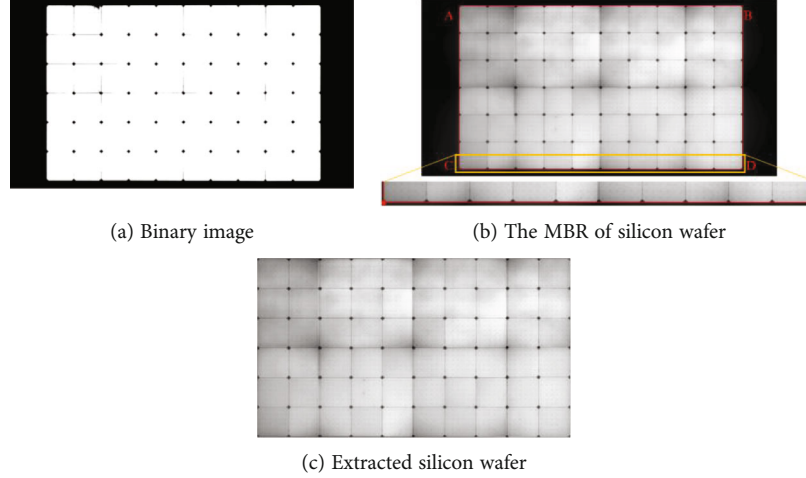
(a) Binary image (b) The MBR of silicon wafer

(c) Extracted silicon wafer

Figure 3: The process of component region extraction with MBR.

Table 1: Description of the corner coordinates of the MBR.

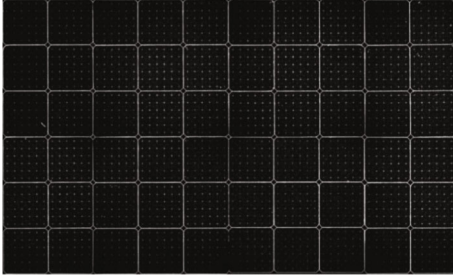| Corner point	A	B	C	D
Coordinate	(x, y)	$(x + w_s, y)$	$(x, y + h_s)$	$(x + w_s, y + h_s)$

Figure 4: The edge binarization image of the silicon wafer region.

binary *Sobel* edge image $f(x, y)$ (with x between 0 and 6333 and y between 0 and 3809, as shown in Figure 4) is projected on horizontal direction and vertical direction, respectively, and then scanned row by row or column by column to make the statistics of validating pixel points. By analyzing the peak value in the projected image, the position information of a single slice in the silicon wafer area image can be determined. The result of projection of binarized edge image is shown in Figure 5.

It can be seen from Figures 5(a) and 5(b) that the vertical and horizontal projection of the image can clearly reflect the distribution of the silicon wafer region and the edge position (as E1 shown in Figures 5(a) and 5(b)). Due to the existence of the laser hole (as shown in Figure 6) and the unevenness of the silicon wafer properties, there are also maintained interference pulses near the actual edges of the silicon wafer (E2 in Figures 5(a) and 5(b)) that will reduce the detection stability of the silicon wafer boundary.

Actually, the interference projection values have certain neighborhood approximation, and the boundary often

appears sudden value. Many methods use adjacent difference to obtain the edge of the target [24, 25]. However, the fixed difference distance is difficult to determine and could weaken the value of the edge itself, especially with gradient. Hence, a local difference extremum of edge projection (LDEEP) is proposed. The main idea is to find the maximum contrast in a neighborhood range for difference, as shown in

$$
\begin{cases}
C'(i) = |C(i) - \min \{C(k) | i - \Delta < k < i\}|, \\
C'(j) = |C(j) - \min \{C(k) | j - \Delta < k < j\}|,
\end{cases} \tag{1}
$$

where C' is the difference between the current projection position and the maximum contrast in the local minimum range, k is the edge index in the neighborhood range, and $\triangle$ represents the local search range.

Figures 5(c) and 5(d) are the LDEEP results obtained by $\triangle$ of 12. It can be seen that most of the interference peaks have been reduced, effectively retaining the edge value of the silicon wafer arrangement. Although the neighborhood difference reduces the absolute peak value of the edge at the E1 silicon wafer, the interference peaks decrease more significantly. That is the contrast between real edge and interference is more obvious and is conducive to obtain the edge position of each column and row of silicon wafers.

2.2. Single Silicon Wafer Segmentation Based on Selective Peak Quick Sort. With the LDEEP results, the first k (one more than the number of silicon wafer in horizontal or vertical direction) positions with maximum values can be looked at as the edge position of the silicon wafer. The bubble sorting algorithm is most commonly used algorithm but needs to search whole projection values, and the time complexity is $O(n^2)$. That is, for a silicon image area with a size of 6333×3809, it needs 6333^2 search times at horizontal direction and 3809^2 search times at vertical direction for bubble sorting algorithm. Actually, for the specific silicon wafer area image (10×6) projection peak order, it is only necessary to obtain the first $k_j (k_j \leq 11)$ boundary values in

(a) Vertical projection

(b) Horizontal projection

(c) Vertical neighborhood difference projection

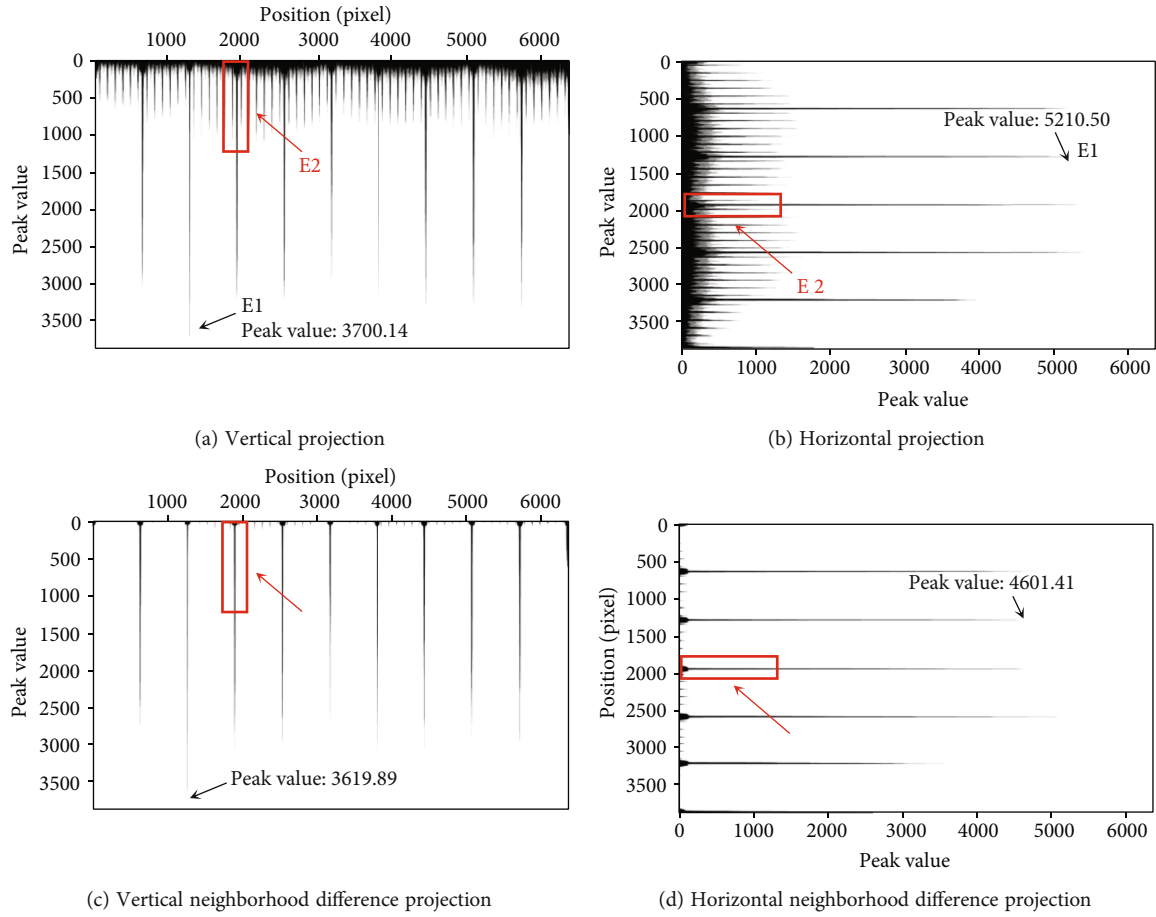(d) Horizontal neighborhood difference projection

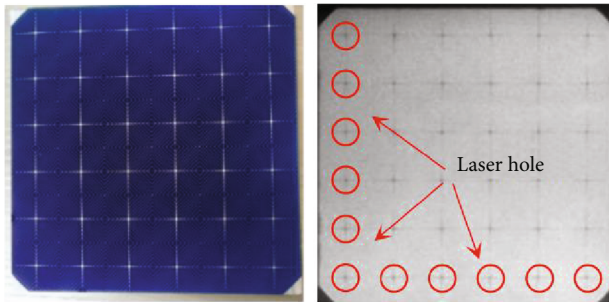FIGURE 5: Projection of edge binarized image with LDEEP of SCC.

FIGURE 6: Front and PL image of single cell.

the horizontal direction, and the first k_i ($k_i \leq 7$) boundary values in the vertical direction, without all of them to be sorted. So a selective peak quick sorting (SPQS) algorithm was presented in this paper. The main idea is to set the distance between two adjacent extreme values. When the distance is less than a threshold, the two extreme values are integrated and the larger one maintained. Figure 7 uses the horizontal projection (from up to bottom in silicon wafer region) as an example for the edge search process of SPQS.

In Figure 7, $\text{Dis}(i, \text{pos}(\mathbf{E}(k)))$ is the distance between the ith position in the neighborhood difference projection vector

$\mathbf{C}'$ and the corresponding position of $\mathbf{E}(k)$, and $\text{End}(\cdot)$ represents the length of the vector. In a traversal process, only finite peak values need to be compared and bubbling, so that the prior conditions (the arrangement of the silicon wafers) can be fully utilized to speed up the search. On the other hand, the value of $\mathbf{C}'$ at the edge of the silicon wafer is often much larger than that at the nonedge position. On the basis of setting the condition T_u (such as $T_u = \max (\mathbf{C}')/2$), the search frequency can be greatly reduced. Table 2 shows the ranking results.

Table 2 shows the search results of edge positions of silicon wafer in a SCC image. According to the results, the boundary location of each silicon wafer is shown in Figure 8. Meanwhile, the single silicon wafer is segmented.

3. Defect Recognition Based on Lightweight Deep Neural Network

When imaging with PL pattern, the laser hole and the surrounding grid lines cannot be excited by ions, so there are black dots and crosses appearance with corner features (as shown in Figure 6). Due to the fragility and stress concentration of the silicon wafer substrate, a kind of internal microcracks (called hidden crack) are prone to occur around the
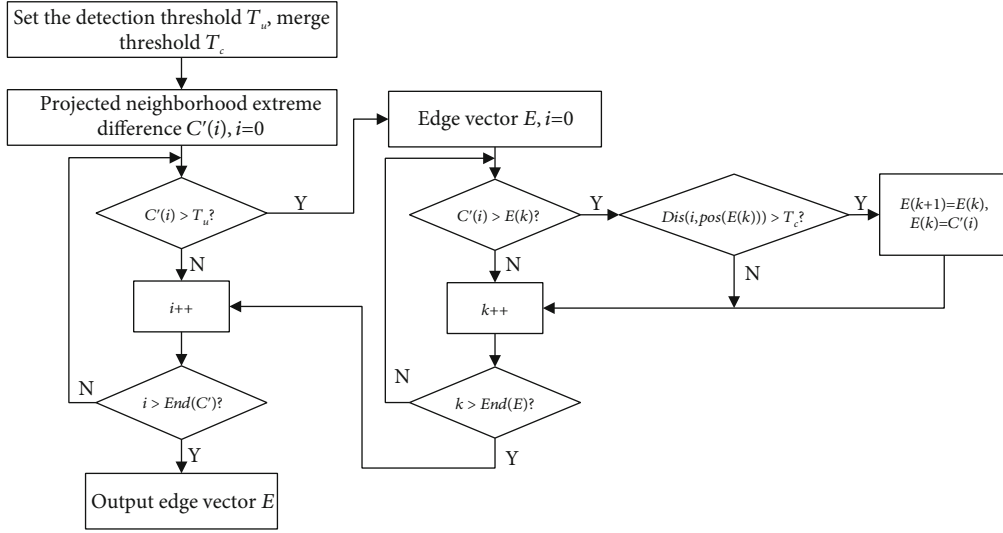
FIGURE 7: Edge search process of SPQS algorithm.

TABLE 2: Positioning results of silicon wafer edge sorting.

Peak number from left to right	Peak coordinates	$C'(j)$ values	Peak number from up to bottom	Peak coordinates	$C'(i)$ values
1	(10.29, 0)	1249.91	1	(0, 12.34)	1372.56
2	(639.03, 0)	2735.72	2	(0, 647.63)	4791.34
3	(1261.22, 0)	3619.89	3	(0, 1289.41)	4601.41
4	(1896.51, 0)	3063.19	4	(0, 1931.31)	4621.05
5	(2525.25, 0)	2958.40	5	(0, 2573.15)	5079.51
6	(3167.09, 0)	2683.32	6	(0, 3221.53)	3560.05
7	(3802.38, 0)	3122.13	7	(0, 3783.73)	1601.79
8	(4437.67, 0)	3036.99			
9	(5066.41, 0)	2879.81			
10	(5714.81, 0)	2945.30			
11	(6326.64, 0)	2080.78			

laser hole (as shown in Figures 9(a) and 9(b)). In the PL image, hidden crack usually presents a darker direction line with 45° diagonals (crack type1) or an "X" shape cross (crack type2) with a size of 0.5 mm~5 mm. Due to the improper detection method in the process of manual inspection, the friction force between the sharp object and the cell is different, resulting in scratches on the surface of the cell, as shown in Figure 9(c). Due to the partial shadow heating in the SCC, a hot spot effect is generated, causing black spots on the cell, as shown in Figure 9(d). These defects are easy to cause damage to the solar cell and seriously affect the life and translate efficiency of the solar cell. Therefore, the detection of cracks, scratches, and black spots is an important content of solar cell detection, and it is also the key and difficult point to realize automatic defect recognition of the SCC.

In order to improve the accuracy of cell detection, this paper provided an improved deep network with the light-weight backbone of YOlOv3-Tiny to detect and recognize the defects. In the traditional YOLOv3-Tiny network structure, the backbone network uses a 3×3 size convolution kernel for feature extraction, and a pooling operation is performed after each feature extraction. Although the network structure is lightweight and has real-time performance, there are still some shortcomings: first, the backbone network convolutional layer has only seven layers, relatively insufficient for detecting hidden cracks with weak and small size. Second, the detection scale information of a network model is less. Only two scales of 13×13 and 26×26 are used to detect single-chip defects, which can easily lead to missed detection of small defects and low detection accuracy.

3.1. Improvement Network Model with Shallow Features. The network structure is deepened with 39 layers and more 1×1 and 3×3 convolutions. The 1×1 convolution kernel is used to adjust the number of channels, 3×3 is used to extract high-dimensional spatial features. Meanwhile, an
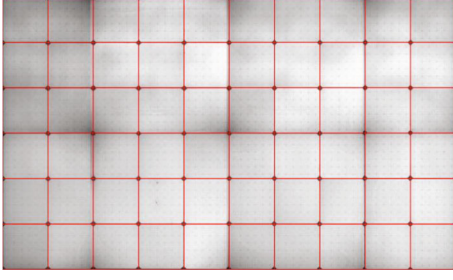
Figure 8: Boundary positioning of silicon wafer arrangement.

upsampling layer on the basis of the original network's 13×13 and 26×26 defects predictions to form 13×13, 26×26, and 52×52 three-scale predictions to further improve the accuracy of defect detection. The improved network structure is shown in Figure 10.

The backbone network consists of 13 convolutional layers and 6 maximum pooling layers. The convolutional layer (Conv) in the network consists of a two-dimensional convolutional layer (Conv2d), a normalized operation layer (BN), and an activation function layer (Leak ReLU) in sequence. The BN layer makes the input of each batch in the effective area of Leak ReLU present a normal distribution, and Leak ReLU can reduce the disappearance of gradients during the training process.

The input image is first converted to 416×416 and becomes the size of 13×13 in the output layer. The output layer has a deeper number of network layers and a larger receptive field, which is suitable for predicting large targets. In the first prediction *Prediction1* on this feature map, the 13×13 feature layer uses a 1×1 convolution kernel to adjust the number of channels, and the output is upsampled. The upsampled feature is stacked with the thirteenth feature layer of the backbone network in the channel dimension, and the output layer with a size of 26×26 is obtained after convolution operation. The output layer has a relatively shallow number of network layers and a small receptive field, which is suitable for small target prediction. *Prediction2* is used for the second prediction on this feature map.

In order to make the proposed model suitable for variant size of defects, the third prediction layer *Prediction3* with shallow information is constructed. In the second prediction path, the 26×26 feature layer uses a 1×1 convolution kernel to adjust the number of channels, and the output is upsampled once, and the upsampled feature is stacked with the ninth feature layer of the backbone network in the channel dimension. After convolution feature fusion, a new output layer with a size of 52×52 is constructed, as shown added prediction scale module in Figure 10. The feature information of the output layer comes from the fusion of the shallow information and the second predicted feature map, which contains more low-dimensional feature information and smaller receptive field, and is suitable for predicting smaller targets. Upon this, the third prediction branch *Prediction3* is designed for different scales fused,

and a certain degree of deep semantic understanding ability can also be provided in the shallow features, so as to give better recognition results of small target objects.

3.2. Prediction of Bounding Box. In the YOLOv3-Tiny algorithm, the picture input into the network is divided into $s \times s$ cells according to the scale of the feature map, and the scale of the feature map is 13×13 and 26×26, respectively. The first prediction layer generates 13×13 grids; the prediction box is larger. The second prediction layer generates 26×26 grids, which is smaller than the prediction box of the first prediction layer.

The crack type1, the crack type2, and black spot are small in the solar cell, and smaller prediction frame is needed. The scale of the third prediction layer constructed is 52×52 in this paper. The input image is finally divided into 52×52 grids. Each grid corresponds to the channel information of the prediction layer. Each prediction layer channel contains the final prediction parameters of the grid.

Taking the first prediction layer as an example, each channel is composed of the offset t_x and t_y of the grid where the center point of the prediction box is located relative to the upper left corner of the grid. $\{t_w, t_h, p_{obj}\}$ are the width, height, and the prediction confidence of the prediction box, $\{s_1, s_2, s_3, s_4\}$ represent the score of the crack type1, the crack type2, the scratch, and black spot, respectively. Each grid in the prediction layer generates three preset bounding boxes according to the corresponding channel information. Each prediction box contains the above 9 parameters, so the dimension of the output channels of the three prediction layers is 27. When training, the bounding box constantly adjusts in size, the predicted box in the grid is matched with the target box in the label, and the bounding box with the largest IOU (Intersection Over Union) value is selected as the output result [26]. The schematic diagram of bounding box prediction is shown in Figure 11.

The formula for obtaining the actual coordinates of the bounding box is as follows:

$$b_x = \sigma(t_x) + c_x, \tag{2}$$

$$b_y = \sigma(t_y) + c_y, \tag{3}$$

$$b_w = p_w e^{t_w}, \tag{4}$$

$$b_h = p_h e^{t_h}, \tag{5}$$

where b_x and b_y are the coordinates of the center point of the bounding box, b_w and b_h are the width and height of the bounding box, c_x and c_y are the coordinates of the upper left corner of the grid where the center of the target box belongs, $\sigma(t_x)$ and $\sigma(t_x)$ are the offset distance of the center coordinates relative to the upper left corner of the grid, and p_w and p_h are the width and height of the preset bounding box.

4. Experimental Results and Analysis

4.1. Sorting Out Single Silicon Wafer Defect Data Set. In this paper, a total of 3200 SCC images with different
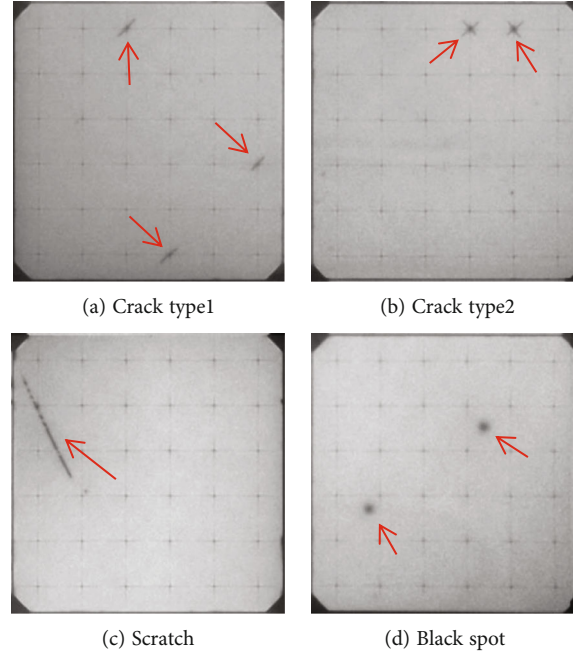
(a) Crack type1

(b) Crack type2

(c) Scratch

(d) Black spot

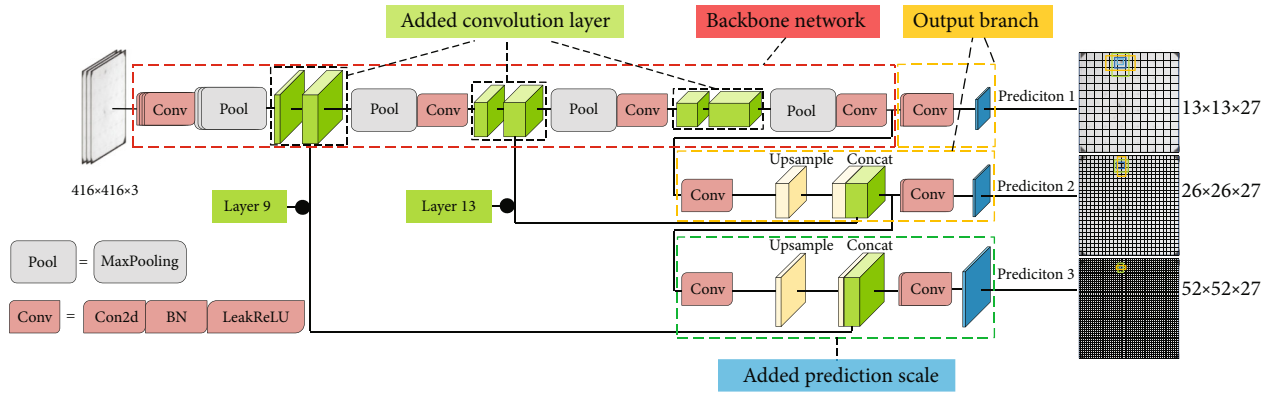FIGURE 9: Types of cell defects detected in this paper.

FIGURE 10: Improved single-cell defect detection network structure.

specifications (10×6 and 12×6) have been collected in a large-scale solar cell factory in Jiangsu province of China. Each silicon wafer in the SCC was segmented and extracted, and 182,400 images of solar cells were obtained. After sample screening, 6100 representative defect samples were selected and maintained. In order to increase the diversity of the data, some defective samples with brightness transformation, flipping, and other operations were added to augment the data set. Finally, 18,000 sample data sets were formed with 4 types of defects: the crack type1, the crack type2, scratch, and black spot. The size of each image is between 630×635 and 665×695 pixels. The defect size in the image is between 64×64 and 85×200 pixels. Part of the data set after data augment is shown in Figure 12.

The defect sample data set was divided into training set, validation set, and test set according to a certain proportion. In order to facilitate the detection of defects, the data set format was converted to the VOC data set format, and the images containing the defects were marked. The number of various data sets after sorting is shown in Table 3.

4.2. Evaluation Index. In order to examine the performance of the improved model, we involved some evaluation indicators including precision (P), recall (R), and mean average precision (mAP). For the VOC data set, if the IOU (the intersection ratio of the predicted box and the real box) is greater than 0.5, the predicted object and the actual object are considered to be the same object; otherwise, it is considered to be not the same object. Under this definition, the calculation formulas for accuracy and recall are

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{6}$$

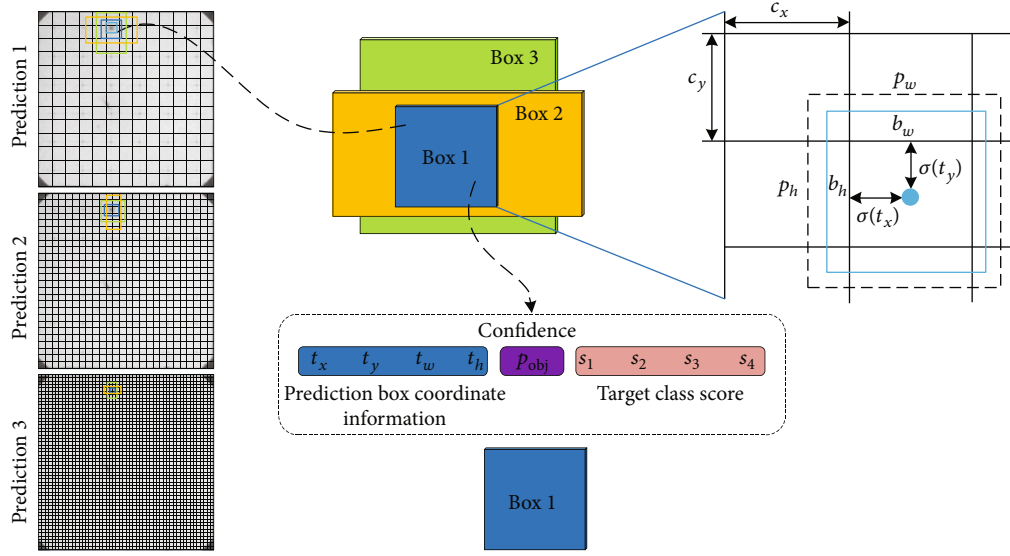$$R = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{7}$$

FIGURE 11: Schematic diagram of the spatial structure of the prediction layer.
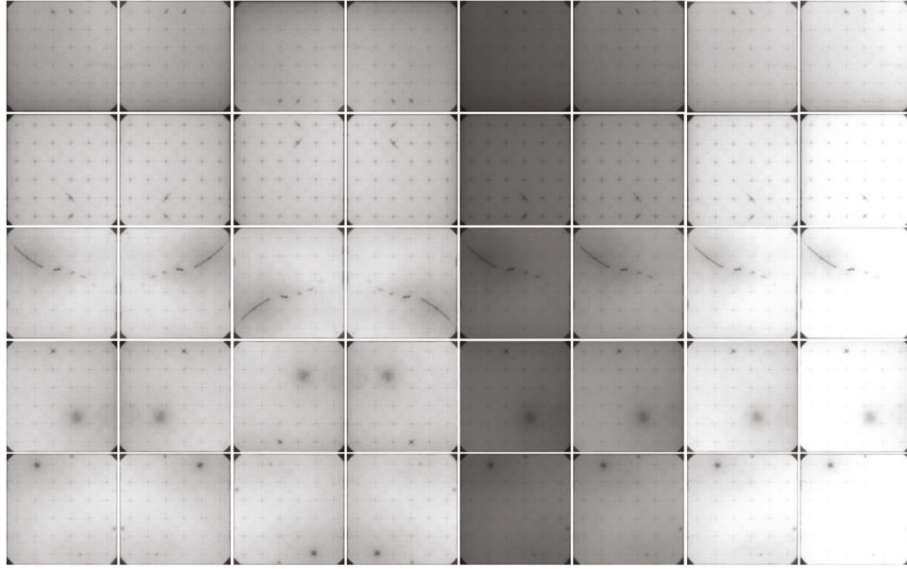

FIGURE 12: Part of the data set after data augment.

TABLE 3: Sorting out the image data set.

Data set	Crack type1	Crack type2	Scratch	Black spot	The total amount
Training set	2800	3800	2400	4000	12000
Validation set	750	800	630	820	3000
Test set	680	820	660	840	3000

where TP represents that a positive sample is predicted to be a positive sample, FP represents that a negative sample is predicted to be a positive sample, and FN represents that a positive sample is predicted to be a negative sample.

With the calculation formulas for precision and recall, a PR curve can be drawn with recall rate as abscissa and precision rate as ordinate for a certain type of target. The area enclosed by the curve is defined as AP. The mAP value can be calculated by the average of the AP values of each class. The calculation formula is as follows:

$$\text{mAP} = \frac{\sum\limits_{k=1}^{n} P(k) \Delta R(k)}{n}. \tag{8}$$

In formula (8), $P(k)$ represents the accuracy rate when k pictures can be recognized, $\Delta R(k)$ represents the change in recall rate when the number of recognized pictures changes from $k-1$ to k, and n is the number of target categories.

(a) Crack type1 PR curve

(b) Crack type2 PR curve

(c) Scratch PR curve

(d) Black spot PR curve
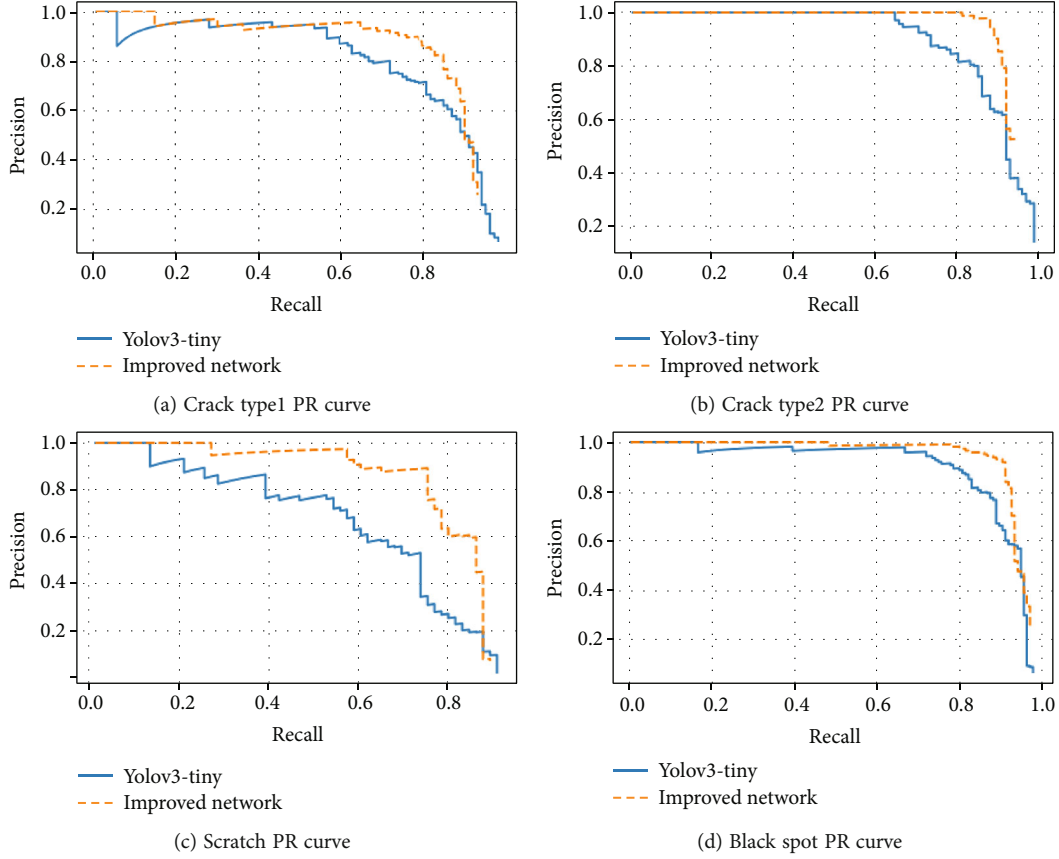
Figure 13: PR curves of YOLOv3-Tiny and improved network.
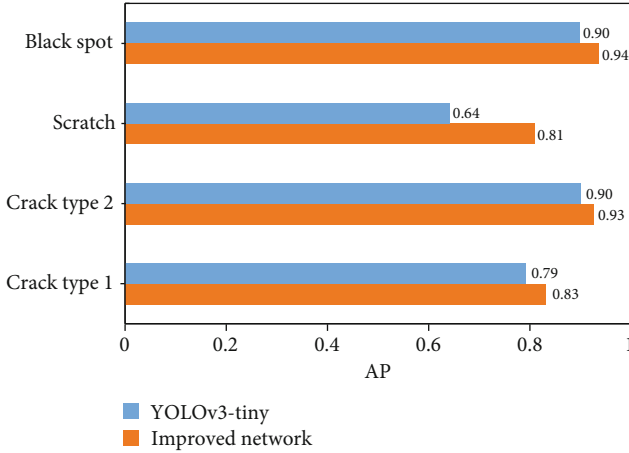
Figure 14: Comparison of AP values of the two network models on different defect types.

Table 4: Performance comparison of different network models.

Network model	mAP (%)	Speed (fps)
SSD	97.52	12
YOLOv3	94.31	25
YOLOv3-Tiny	80.77	47
Improved network	87.55	40

4.3. Experimental Results and Analysis.

The experiment was carried out under Windows operating system, using the Darknet deep learning framework [27], with NVIDIA GTX 1050, Intel (R) core i5-7300HQ, 8 G memory, install CUDA10.0, cudnn7.4.1 to achieve GPU acceleration, using Python language to achieve the improved algorithm. Set the maximum number of iterations for model training to 20,000, the initial learning rate is 0.001, after 10,000 itera-tions, the learning rate is 0.0001, the batch is 64, and the sub-batch is 8.

The PR curves of YOLOv3-Tiny and improved algo-rithm for different types of defects are shown in Figure 13. The areas under the PR curve in the figures are the AP value of the corresponding class. The larger the area under the curve, the higher the corresponding AP value and the better detection performance.

The comparison of AP values between YOLOv3-Tiny and the improved network algorithm on the crack type1, the crack type2, scratch, and black spot is shown in Figure 14.

From Figures 13 and 14, it can be seen from the figures that the overall AP value of scratch is the lowest among all types. That is because the training data set of scratches is thin and small, it is prone to overfitting, and the generaliza-tion ability on the test set is poor. On the other hand, as the
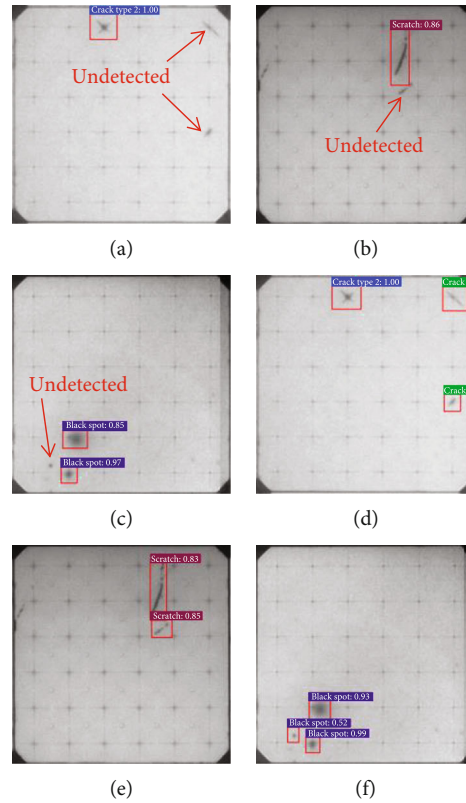
FIGURE 15: Comparison of test results before and after the model improvement: (a), (b), and (c) show the effect before the improvement; (d), (e), and (f) show the effect after the improvement.

diversified shapes of scratch defects, false detections are prone to occur during the testing process.

Furthermore, the scratch AP value reaches the most absolute increase with 0.17, which indicates the detection accuracy of defect target improved by multiscale feature fusion layers. The AP value of the crack type2 is the least increase with only 0.03. This is mainly because images of crack type2 have more training data sets, single defect shape, and better detection effect. From the overall effect, the improved algorithm has better detection performance and better generalization ability.

Generally, the experiment proved that increasing the depth of the network and the number of convolution kernels could improve the feature extraction capability; meanwhile, it may take more time consuming. In order to inspect the real-time performance of the improved algorithm, the mAP value and speed are analyzed. Under the same experimental conditions, SSD, YOLOv3, YOLOv3-Tiny, and the improved network model are tested on the test data set. The test results are shown in Table 4.

As shown in Table 4, the mAP values of SSD and YOLOv3 are 97.52% and 94.31%, respectively, which are much higher than those of YOLOv3-Tiny and the improved algorithm. However, due to the large network model, many parameters bringing large amount of calculation, the detection speed is slow [28], which is difficult to adapt the real-time detection of solar cell on the production line. The mAP value of the improved network model is 87.55% with 6.78% higher than that of the YOLOv3-Tiny network model, indicating that the detection performance of the improved network model has improved. Because the improved algorithm deepens the network structure and increases the computation of the network model to a certain extent, the detection speed of the improved network model is slightly lower than that of YOLOv3-Tiny. Generally, the detection speed above 25 fps can meet the real-time requirements of target detection, and the improved algorithm can still meet the real-time requirements. It can be seen that the improved network model can take into account both the detection speed and accuracy, which can better complete the task of solar cell defect detection.

In order to show the detection effect of the improved network model more intuitively, Figure 15 shows the comparison of the detection results before and after the model improvement.

In order to verify the feasibility of the improved model in this paper, various types of component images have been collected on the SCC production line. These images are randomly collected at different times, locations, and under different lighting conditions. Take the 10×6 specification module as an example; the SCC is composed of 60 cell silicon wafers arranged. The improved model detection speed is 40 fps, and the average time per cell is 0.025 s. Therefore, the average time to complete a SCC detection is 1.5 s, which meets the real-time detection requirements. The effect of using the improved network model to detect different types
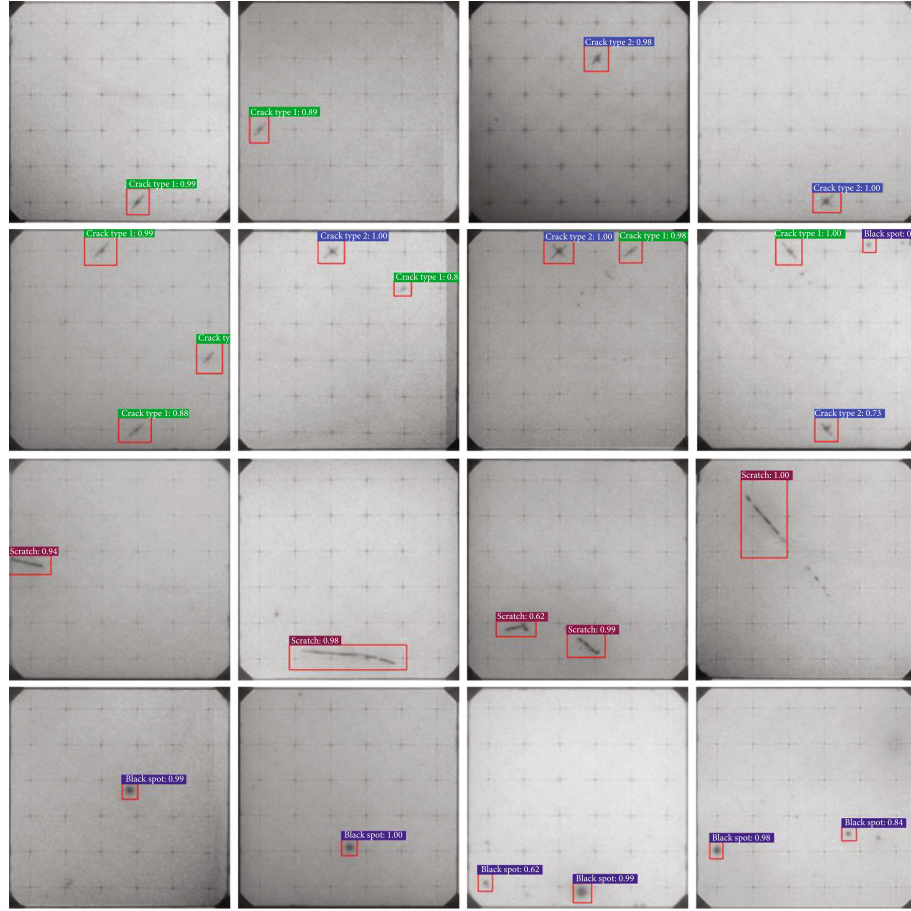
FIGURE 16: The effect of improved defect detection.

5. Conclusion

In the SCC PL image, the defect area accounts for less than one ten thousandth, which brings difficulties to the efficient detection and recognition of defects. On the premise of ensuring the efficiency, in order to improve the detection accuracy, the SCC units are segmented and extracted, which lays the foundation for the next step of defect detection and recognition. In order to recognize the weak and small defects, an improved network model based on YOLOv3-Tiny is proposed. We deepened the structure of the main feature extraction network, added a feature pyramid layer to realize three-scale prediction, and strengthened the semantic information of the shallow feature map. On the solar cell defect test data set, the recognition rate of mAP is 87.55%, which is 6.78% higher than the original algorithm, and the detect speed is 40 fps, which meets the requirements of real-time detection. The experimental results show that the improved model improves the detection accuracy, reduces the missed detection rate, and achieves a good balance between detection accuracy and detection speed.
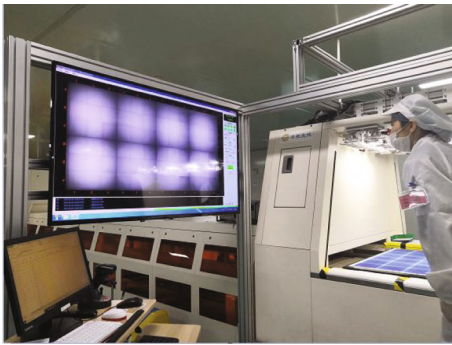

FIGURE 17: Solar cell component detection site.

of solar cell defects is shown in Figure 16. It can be seen that the improved network model can more accurately detect the crack type1, the crack type2, scratch, and black spot.

Now, the detection model was successfully applied in the field and worked well (as shown in Figure 17). On-site operation shows that the defect detection method can ensure the speed and accuracy of SCC defect detection and lay the foundation for automatic identification of defects in SCC products.

Whereas, due to the limited model parameters, when the environment is more complicated (such as occlusion and dark light), the detection accuracy of the model still needs to be improved. On the premise of both speed and accuracy,

improving the generalization ability of the model will be the main research direction in the future.

Data Availability

The [DATA TYPE] data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] E. Kabir, P. Kumar, S. Kumar, A. A. Adelodun, and K. H. Kim, "Solar energy: potential and future prospects," *Renewable & Sustainable Energy Reviews*, vol. 82, pp. 894–900, 2018.

[2] M. Dhimish, V. Holmes, B. Mehrdadi, and M. Dales, "The impact of cracks on photovoltaic power performance," *Journal of Science: Advanced Materials and Devices*, vol. 2, no. 2, pp. 199–209, 2017.

[3] X. L. Qian, H. Q. Zhang, Y. X. Chen et al., "Research development and prospect of solar cells surface defects detection based on machine vision," *Journal of Beijing University of Technology*, vol. 43, no. 1, pp. 76–85, 2017.

[4] S. Deitsch, V. Christlein, S. Berger et al., "Automatic classification of defective photovoltaic module cells in electroluminescence images," *Solar Energy*, vol. 185, pp. 455–468, 2019.

[5] L. Zhang, P. Liang, H. S. Zhu, and P. D. Han, "Detection of finger interruptions in silicon solar cells using photoluminescence imaging," *Chinese Physics B*, vol. 27, no. 6, article 068801, 2018.

[6] H. Chen, H. Zhao, D. Han, W. Liu, P. Chen, and K. Liu, "Structure-aware-based crack defect detection for multicrystalline solar cells," *Measurement*, vol. 151, article 107170, 2020.

[7] D. M. Tsai, G. N. Li, W. C. Li, and W. Y. Chiu, "Defect detection in multi-crystal solar cells using clustering with uniformity measures," *Advanced Engineering Informatics*, vol. 29, no. 3, pp. 419–430, 2015.

[8] H. Chen, Y. Pang, Q. Hu, and K. Liu, "Solar cell surface defect inspection based on multispectral convolutional neural network," *Journal of Intelligent Manufacturing*, vol. 31, no. 2, pp. 453–468, 2020.

[9] L. Liu, C. Wang, and S. W. Zhao, "Research on solar cells defect detection technology based on machine vision," *Journal of Electronic Measurement and Instrumentation*, vol. 32, no. 10, pp. 47–52, 2018.

[10] X. B. Wang, J. Li, M. H. Yao, W. X. He, and Y. T. Qian, "Solar cells surface defects detection based on deep learning," *Pattern Recognition and Artificial Intelligence*, vol. 27, no. 6, pp. 517–523, 2014.

[11] L. Wu, M. Liu, Q. Jiang, M. Ge, and L. Ling, "Solar cell surface quality detection system based on artificial neural network," *Journal of Hefei University of Technology (Natural Science)*, vol. 9, 2017.

[12] C. Chen and Q. I. Feng, "Review on development of convolutional neural network and its application in computer vision," *Computer Science*, vol. 46, no. 3, pp. 63–73, 2019.

[13] C. L. Zitnick and P. Dollar, "Edge boxes: locating object proposals from edges," in *European Conference on Computer Vision*, pp. 391–405, Zurich, 2014.

[14] 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*, pp. 580–587, Columbus, OH, USA, 2014.

[15] R. Girshick, "Fast R-CNN," in *Proceedings of the IEEE international conference on computer vision*, pp. 1440–1448, Santiago, Chile, 2015.

[16] 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 & Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017.

[17] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6517–6525, Las Vegas, NV, USA. New York, 2016.

[18] W. Liu, D. Anguelov, D. Erhan et al., "SSD: single shot multibox detector," *Champions*, vol. 9905, pp. 21–37, 2016.

[19] J. Li, Z. Su, J. Geng, and Y. Yin, "Real-time detection of steel strip surface defects based on improved YOLO detection network," *IFAC-PapersOnLine*, vol. 51, no. 21, pp. 76–81, 2018.

[20] Z. Qiu, S. Wang, Z. Zeng, and D. Yu, "Automatic visual defects inspection of wind turbine blades via YOLO-based small object detection approach," *Journal of electronic imaging*, vol. 28, no. 4, pp. 43023.1–43023.11, 2019.

[21] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *IEEE Conference on Computer Vision & Pattern Recognition*, pp. 6517–6525, Honolulu, HI, USA, 2017.

[22] H. Gong, H. Li, K. Xu, and Y. Zhang, "Object detection based on improved YOLOv3-tiny," in *2019 Chinese Automation Congress (CAC)*, pp. 3240–3245, Hangzhou, China, China, 2019.

[23] Z. Yi, S. Yongliang, and Z. Jun, "An improved tiny-yolov3 pedestrian detection algorithm," *Optik*, vol. 183, pp. 17–23, 2019.

[24] Q. C. Sun, Y. Q. Hou, and Q. C. Tan, "A subpixel edge detection method based on an arctangent edge model," *Optik*, vol. 127, no. 14, pp. 5702–5710, 2016.

[25] X. Liu and F. Xue, "Moving object detection based on edge contrast difference algorithm," *Computer Engineering*, vol. 44, no. 10, pp. 246–251, 2018.

[26] B. Jiang, R. Luo, J. Mao, T. Xiao, and Y. Jiang, "Acquisition of localization confidence for accurate object detection," in *Proceedings of the European conference on computer vision (ECCV)*, pp. 784–799, Munich, Germany, 2018.

[27] R. Guo, S. Li, and K. Wang, "Research on YOLOv3 algorithm based on darknet framework," *Journal of Physics: Conference Series*, vol. 1629, article 012062, 2020.

[28] J. H. Park, H. W. Hwang, J. H. Moon et al., "Automated identification of cephalometric landmarks: part 1—comparisons between the latest deep-learning methods YOLOV3 and SSD," *The Angle Orthodontist*, vol. 89, no. 6, pp. 903–909, 2019.