Detection of surface defects on solar cells by fusing Multi-channel convolution neural networks

Xiong Zhang[a,*], Yawen Hao[a], Hong Shangguan[a,b], Pengcheng Zhang[b], Anhong Wang[a]

[a] *Shanxi Key Laboratory of Advanced Control and Equipment Intelligence, Taiyuan University of Science and Technology, Taiyuan 030024, China*
[b] *Shanxi Provincial Key Laboratory for Biomedical Imaging and Big Data, North University of China, Taiyuan 030051, China*

A B S T R A C T

Manufacturing process defects or artificial operation mistakes may lead to solar cells having surface cracks, over welding, black edges, unsoldered areas, and other minor defects on their surfaces. These defects will reduce the efficiency of solar cells or even make them completely useless. In this paper, a detection algorithm of surface defects on solar cells is proposed by fusing multi-channel convolution neural networks. The detection results from two different convolution neural networks, i.e., Faster R-CNN and R-FCN, are combined to improve detection precision and position accuracy. In addition, according to the inherent characteristics of the surface defects in solar cells, two other strategies are used to further improve the detection performance. First, the anchor points of the region proposal network (RPN) are set by adding multi-scale and multi-aspect regions to overcome the problem of high false negative rate caused by the limitation of anchor points. Second, in view of the subtle and concealed defects of solar cells, the hard negative sample mining strategy is used to solve the problem of low detection precision caused by the negative sample space being too large. The experimental results showed that the proposed method effectively reduced the false negative rate and the false positive rate of a single network, and it greatly improved the accuracy of the locations of defects while improving the object recall rate.

1. Introduction

Facing the increasing crisis of serious ecological and environmental damage as well as the traditional energy shortages, people have identified some new sources of energy, such as solar energy, which is environmentally friendly and non-polluting [1]. At present, solar energy is collected mainly by solar cells, which are photoelectric conversion carriers that are based largely on crystalline silicon. The quality of the silicon wafer directly determines the conversion efficiency of solar cells. An electroluminescence (EL) flaw detector shown in Fig. 1 is used to judge whether solar cells are defective [2]. The solar cell is excited by a forward current of 1–40 mA, and the atoms in an excited state are unstable. Subsequently, a cooled Si-CCD camera is used to capture infrared light that is emitted spontaneously. Through the relationships between the lifetime of the minority carrier, density, and light intensity, it is judged whether there is a defect in the silicon wafer based on the extent of the exposure of the negative film [3]. However, the defects in production technology and artificial misoperation often lead to subtle defects on the surface of solar cells, such as broken cells, cracks, over-welding/broken gates, black edges, black spots, and short circuits. It is

of great theoretical significance and practical value to study the surface defect detection technology of solar cell module, and then effectively eliminate the defective solar cells.

At present, there are three main types of methods for detecting defects in solar cell EL images, i.e., visual inspection, physical methods, and machine vision. Among them, the defect detection method based on machine vision has been greatly developed due to its advantages of good real-time performance, high accuracy, and convenient operation. From the perspective of mathematical modeling, detecting surface defects can be divided into three types of methods, i.e., the image domain analysis methods (e.g., the gradient feature method [4,5], the matrix decomposition method [6] and the clustering method [7]), the transform domain analysis method (e.g., the Fourier transform method [8] and the wavelet transform method [9]), and the deep learning method [10,11]. The traditional image domain and transform domain methods extract features mainly by artificially designed filters, and have weak ability to capture high-level semantic features. The traditional algorithms can work well in the situations of rule defects and simple scenes, but it is no longer applicable for the situations where the features are not obvious and the shapes of defects are diverse and chaotic.
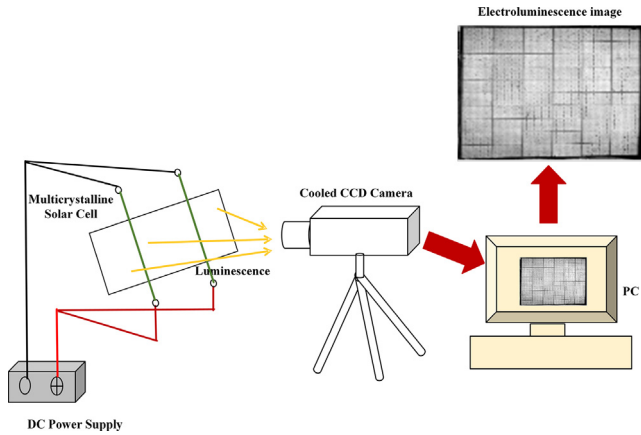
Fig. 1. EL imaging system.

Compared with these traditional algorithms, deep learning could extract deep abstract features due to a powerful statistical learning ability, and have stronger robustness and better generalization ability when applied to real application environments. In the training process of deep learning, there is no need for complex adjustment and human intervention.

In 2006, Geoffrey Hinton [12] first proposed the concept of deep learning. The first deep learning model is deep convolutional neural networks (CNN) that can automatically learn the features of samples, in which low-dimensional shallow features are mapped to high-dimensional high-level features. Thus, they have significant performance advantages in recognizing images, classifying objects, generating text, generating the content of images, annotating images, tracking video objects, machine translation, controlling robots, and other fields. In 2014, Ross Girshick [13] proposed the Region Proposal-based object detection framework R-CNN to solve the regional location problem. Then, he proposed Fast R-CNN [14], which used the loss functions of multiple tasks and updated all layers during training without features saved. Fast R-CNN is an accelerated version of RCNN. Although it solves the time-consuming training problems of R-CNN, i.e., slow test speed and repeated calculations, it still has low detection efficiency and cannot meet the requirement of real-time applications. In 2015, Ross Girshick's team proposed Faster R-CNN [15] by replacing the original selective search method of Fast R-CNN with the Region Proposal Network (RPN) to generate candidate bounding boxes. Faster R-CNN greatly improved the detection speed compared with Fast R-CNN. In 2016, the region-based, fully-convolutional networks (R-FCN) were proposed by Dai J [16] in response to the problem that the translation variability of convolutional neural networks becomes worse as the depth increases. R-FCN moved the convolution before the Region of Interest (ROI) layer. R-FCN can improve the object detection rate while maintaining high-accuracy positioning by utilizing the feature maps that are more sensitive to the location.

Inspired by the successful application of deep learning, more and more methods [17,18] use deep learning to detect surface defects. This kind of methods use deep neural network to obtain the mapping relationship between training samples and defect-free samples. However, the complex process of selecting candidate regions often leads to low detection speed and low detection accuracy. In this work, we use an anchor based network instead of semantic information based network to detect the surface defects on solar cells, mainly considering the following factors: firstly, on the surface of solar cell, the defect size is much smaller than the whole picture, only accounting for 0.1% to 1% of it, but the whole picture has a minimum size of 5232 × 2720, therefore pixel by pixel detection will cause huge time consumption; secondly, in practical application, the defective solar cells will be replaced directly, so the repair process does not require the specific shape of the defect. Although accurate defect shapes can be obtained by using semantic

information networks, it comes at the expense of processing speed.

In this paper, two basic detection frameworks are integrated, i.e., Faster R-CNN based on VGG-19 and full convolution R-FCN based on Resnet-101. The proposed method increases the number of anchor points in the RPN, and adds the different scales and ratios of the bounding box in the training process. In addition, the strategy of hard negative sample mining was used with the aim of identifying the subtle hidden defects of solar cells. The outputs of the two networks were complementarily fused to improve the performance of the detectors. The experimental results showed that this method effectively overcame the high false negative and false positive rates of a single network, and that it provided significant improvements in the object recall rate and the accuracy of the defect location regression.

2. Background

The essence of detection problem is to classify and locate the objects to be detected accurately. In order to detect defects on the surface of solar cells, two problems must be solved, i.e., different defects must be classified and located (bounding box regression). A brief review of two neural network models based on deep learning related to this study is provided in this section.

2.1. Faster R-CNN

Object detection includes three basic steps, i.e., features extraction, candidate RoI region generation, and object classification and location regression. Faster R-CNN covers these three basic steps in the framework of one deep neural network and achieves end-to-end detection of the object. Fig. 2 shows the structure of Faster R-CNN detection.

Specifically, the Faster R-CNN detection model consists of two modules, i.e., the RPN network (region proposals module) and the Fast R-CNN network (the detection module). First, an image of any size is used as input to the detection network. Second, the convolution layer is used to extract features. Third, the RPN network is used to generate high-quality region proposals, and these region proposals are mapped to the last layer where the convolution feature map of CNN is completed. Fourth, the RoI pooling layer is used to fix the size of each region proposals. Fifth, the box-classification layer and box-regression layer are used to make a specific classification judgment and accurate border regression for the proposed region. When the test is being conducted, Fast R-CNN shares the convolutional layer with the RPN. Fig. 3 shows the structure of a RPN network. Specifically, the RPN network traverses the feature map of the last convolutional layer by a sliding window, and it inputs the low-dimensional short vectors into two parallel, fully-connected network layers, i.e., the EL image defect regression layer (reg) and the EL image defect classification layer (cls).

2.2. R-FcN

Unlike classification problems, detecting defects requires that the object features respond accurately to the translation of the object in the
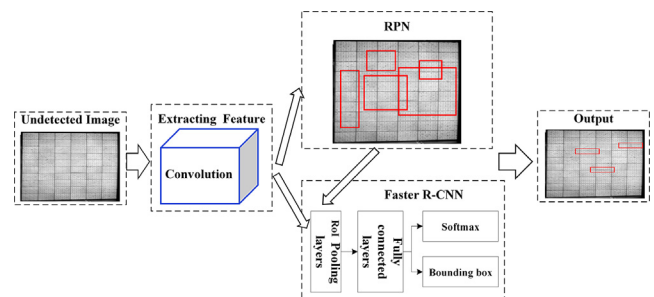

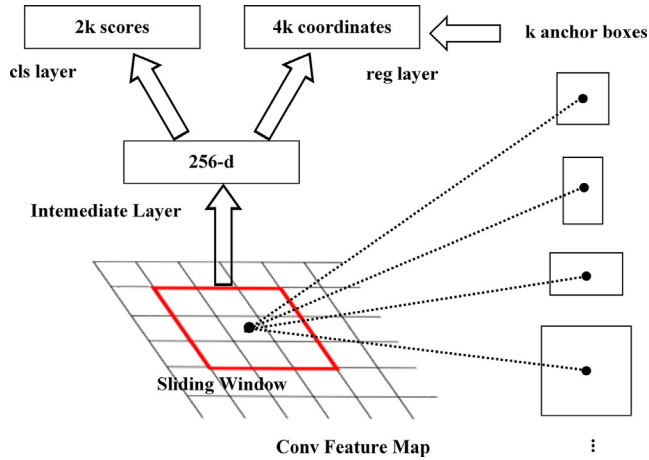**Fig. 2.** Structure of Faster R-CNN.
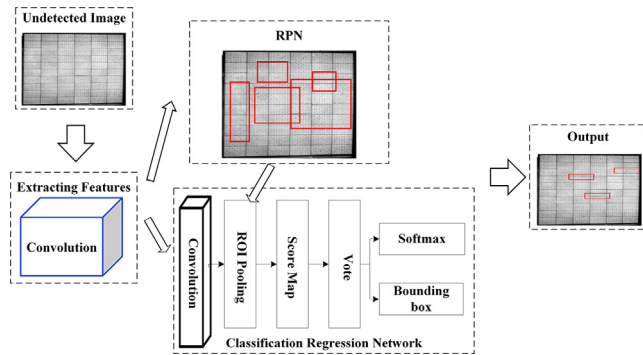
Fig. 3. Structure of RPN network.

Fig. 4. R-FCN network structure.

Fig. 5. Several common surface defects in EL images of solar cells.

images. As shown in Fig. 4, in the R-FCN framework, all the time-consuming convolution layers are placed on the front shared sub-network, and only one layer of convolution operation is needed to make predictive judgments. This design of the network structure greatly reduces the amount of computation and overcomes the shortcomings of Faster R-CNN in translation variability and computational speed. R-FCN removes the fully connected layer in the framework of classical, region-extraction-based R-CNN series networks (e.g., SPP Net [19], Fast R-CNN, Faster R-CNN) and places all of the time-consuming convolution layers that are used to extract the features of images on shared sub-networks. R-FCN uses the last convolutional layer to calculate the position-sensitive RoI score map, which can fuse the position information of the object into the RoI pooling layer. FCN is an improved network of Faster R-CNN. The overall detection process of R-FCN is as follows: ① Input images of any size into the shared CNN for the extraction of features through a series of convolution operations; ② Use the RPN network to generate high-quality suggestion bounding boxes (each picture generally generates approximately 300 suggestion bounding boxes); ③ Calculate the position-sensitive RoI score map for each suggestion bounding box using the last convolution operation layer; and ④ Vote for each position score map, and perform softmax classification judgment and bounding box regression.

3. Method

Due to the influence of the production process and manual operation, the collected EL images often contain a certain amount of noise. Fig. 5 shows that defects, such as cracks, broken cells, and unsoldered cells, are very subtle and concealed. In some areas, the gray value difference between the background and the solar cell is very small. Therefore, the detection of defects in solar cells is a challenging task.
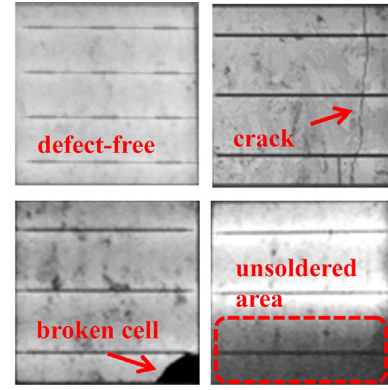
In recent years, deep convolutional neural networks have been used extensively in detecting objects. Utilizing its unique RPN Network, Faster R-CNN could well take into account defects in different positions and detect defects accurately in different locations. However, the different types of defects on solar cells have different shapes and contours. For example, cracks generally have somewhat elongated shapes, and unsoldered cells are somewhat square in shape, and their scales vary significantly. Restricted by the number of anchors in the RPN network, the deviation between the object bounding box and the real bounding box is large, so the accuracy of the position of the detection box of Faster RCNN is reduced. On the basis of Faster R-CNN, R-FCN uses the complete convolution network to detect defects, and it adds a position-sensitive score map to the network. Due to these improvements, the regression accuracy of candidate boxes in R-FCN networks has been improved greatly. However, the position-sensitive score map decomposes each position in the network, resulting in R-FCN having a higher false negative rate than Faster R-CNN.

In this research, in order to improve the precision of the detection position, reduce the false negative rate and the false positive rate, we designed the convolutional neural network structure by fusing the Faster R-CNN and the R-FCN network. The specific idea is as follows:

Multi-Channel Fusion: We proposed a multi-channel fusion scheme by combining the detection results from R-FCN and Faster R-CNN, to improve the regression accuracy of the defect position significantly and reduce the false negative rate and the false positive rate that occur in a single detection model.

Add Multi-scales and Multi-ratios: Considering the Faster RCNN detector is limited by the number of anchors in the RPN network, and the deviation between the object bounding box and the real bounding box is large, so we improved the RPN network by adding multi-scales and multi-ratios, which makes the regression box more precise and greatly reduces the false negative rate.

Online Hard Examples Mining(OHEM): Since the defect area to be detected in the defective image of solar cells is much smaller than the background area, and the large negative sample space makes it difficult to converge in training. Utilizing the OHEM strategy can enhance the discriminant ability and achieve desirable defects detection performance.

The main task of this work is to detect three types of defects (cracks, broken cells, and unsoldered cells). The EL images of solar cells identify all of the defects in each EL image, and they locate the different defects accurately (give the bounding box). The whole detection process of this study was divided into four parts, i.e., (1) EL image defects dataset construction, (2) network training, (3) multi-channel fusion, and (4) defective EL images testing. Fig. 6 shows the overall detection framework.
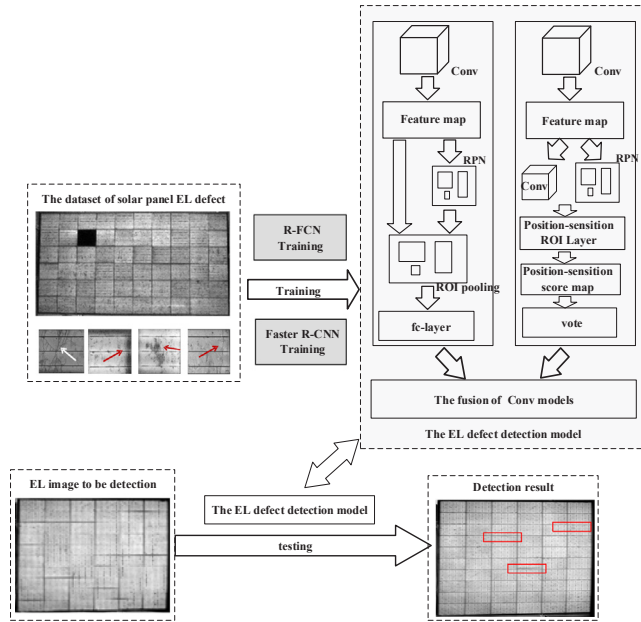
Fig. 6. Illustration of the proposed Multi-channel fusion model for EL image defect detection.

3.1. Network training

Due to the limitation of the anchor number, the RPN network in the original Faster R-CNN model was prone to fail to detect small object defects in EL images and provide inaccurate locations. Considering that the cracks always are slender and long, and the sizes of the broken cells are always variable, in this study, we added two new scales (32,64) and two new aspect ratios (1:3, 3:1) to the RPN network, and the final anchor numbers were increased from 9 to 25. These changes enabled our trained models to learn features in various sizes, and the candidate frames extracted by RPN are more accurate and more applicable. Table 1 shows the anchor parameters of the RPN network in this research.

In the process of training EL images, a big problem is that the object defect area in the entire image is much smaller than the background area, and the negative sample space is large. If such extremely unbalanced data were taken directly to the training model, the model would tend to divide all of the samples into negative samples, which makes it difficult to converge in training. The problem of dataset imbalance, i.e., the number of negative samples is far greater than the number of positive samples, can be solved by the OHEM strategy. Its core idea is to screen out hard examples, i.e., the samples that have great influence on classification and detection, according to the loss function of the RoIs of input samples, and then apply the selected hard samples to the training of the descent of the stochastic gradient.

The detection network is divided into ConvNet and RoINet. The ConvNet is shared as the underlying convolution layer; RoINet includes the RoI Pooling layer, several full connection layers, and two loss layers. ConvNet was used to get the feature maps, which is the input for the RoI Pooling layer. All of the obtained feature maps were fed into

Table 1
Parameter setting of RPN network's anchors.

	Anchor	Number
Base_Size	256 × 256	1
Original Ratios	[1:1, 1:2, 2:1]	3
Original Scale	[128, 256, 512]	3
New Ratios	[1:1, 1:2, 1:3, 3:1, 2:1]	5
New Scale	[32, 64, 128, 256, 512]	5

convolutional neural networks as the input to the RoI pooling layer. Pre-computed proposals were projected onto feature maps through the RoI Pooling layer, and a fixed-feature output was obtained as the input of the full connection layer.

In order to reduce the use of memory and the time required for backward propagation, two ROI networks sharing weights were used. The first ROI Network had only a forward operation, in which the loss of the proposed Rols is calculated, and non-maximum suppression (NMS) was used to merge the candidate ROI boxes. According to the loss ranking from low to high, the first B candidate ROI boxes with lower score were selected from the N candidate boxes as the hard samples. The latter ROI Network included forward and backward operations, and it used the selected hard samples as input to calculate the loss and return the gradient to update the parameters of the entire network. Only the hard samples can update the parameters by back propagation, so there will be no redundant back propagation calculation.

The training stage mainly included data processing, network training, and parameter optimization. In the testing stage, the effectiveness of the trained model was tested with test samples. Fig. 7 shows the specific training-test process. Faster R-CNN uses an alternative training method to alternately train the RPN network and the Fast R-CNN network. R-FCN uses an end-to-end training mode that does not need to divide multiple stages to train the network, and this makes the network training process simple and easy to operate. Table 2 provides the basic parameters of the network settings during training.

3.2. Multi-Channel fusion scheme

In this paper, Faster R-CNN and R-FCN are used as the basic detection framework, and the detection results of the two networks are fused to improve the performance of the detector. The following three reasons explain why we fuse the detection results of those two convolutional neural networks.

(a) R-FCN's unique position-sensitive score map has a natural advantage for the regression of the defect bounding box, which can compensate for the low accuracy of Faster RCNN very well.

(b) In order to reduce network complexity and improve detection speed, R-FCN puts all convolutional layers in front of position-sensitive RoI pooling layer, increasing the number of layers sharing parameters. Doing so will break the translation invariance of the original convolutional network. Although R-FCN adopts the position-sensitive score map to make up for the lack of translation invariance, it still has poor detection accuracy for surface defects with translation changes. Fusing Faster R-CNN with R-FCN can further compensate for the poor translation invariance of R-FCN network, thus improving the detection accuracy of the whole network.

(c) The characteristics of the network structures of Faster R-CNN and R-FCN make the false negative rate and the false positive rate unbalanced. Faster R-CNN has a high false positive rate, and R-FCN has a low false-positive rate. We used the fusion strategy to deal with the problem of unbalanced detection performance caused by a single network.

Fig. 8 shows the schematic diagram of fusion scheme. In our research, the category score and the coordinates of the regression box appeared in R-FCN are represented by Rs (R-FCN score) and Rb (R-FCN bounding box), respectively. The category score and the detected position coordinates that appeared in Faster R-CNN are represented by Fs (Faster R-CNN score) and Fb (Faster R-CNN bounding box), respectively. The category score and regression box position coordinates that appeared after fusion are represented by Bs and Bb, respectively.

The specific steps of fusion are as follows:

Step 1: Determine whether both R-FCN and Faster R-CNN have
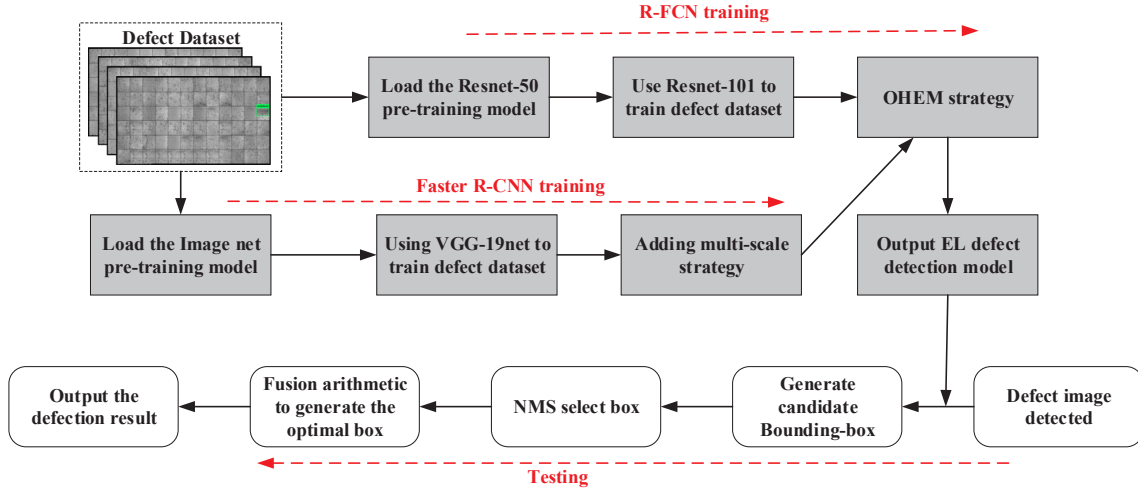
Fig. 7. Flowchart for Training and Testing: R-FCN Training uses the Resnet model with the OHEM strategy; Faster R-CNN Training uses the VGG-19 Net, and adds the Multi-scale (aspect-ratio) and OHEM strategy.

Table 2
Basic parameter setting for Network-training.

	Faster R-CNN	R-FCN
Batch_size	64	128
Base lr	0.001	0.001
Lr-adjustment factor $\begin{cases} s_1 = R_s/(R_s + F_s) \\ s_2 = F_s/(R_s + F_s) \end{cases}$	0.1	0.1
Momentum $\begin{cases} R_x = R_b(x) + R_b(w)/2 \\ R_y = R_b(y) + R_b(h)/2 \\ F_x = F_b(x) + F_b(w)/2 \\ F_y = F_b(y) + F_b(h)/2 \end{cases}$	0.95	0.9
Weight decay(R_x, R_y)	0.0004	0.0005
Iterations N	120,000	120,000

output bounding box.

(Step 2.a) If both of the IOU values are greater than 0, the score of the output bounding box Bs will increase:

$$B_s = R_s \times Weight_{R_2} + F_s \times Weight_{F_2}$$
$$1 < Weight_{R_2} + Weight_{F_2} < 2 \qquad (2)$$

where the weight is determined in two cases: If $R_s > F_s$, assign the value of $Weight_{F_1}$ to 1, and then $Weight_{F_1} = 1 - R_s$, thus $B_s = R_s + F_s \times (1 - R_s)$; If $R_s \leqslant F_s$, assign the value of $Weight_{F_1}$ to 1, and then $Weight_{R_1} = 1 - F_s$, thus $B_s = R_s(F_s - 1) + F_s$.

The addition of the position-sensitive score map in R-FCN enables the detection to respond accurately to the translation of the object. Therefore, when both R-FCN and Faster R-CNN have output bounding boxes, the position of the fusion prediction bounding boxes retains Rb.

(Step 2.b) If both of the two IOU values are equal to 0, suppress the score of the output bounding box; the suppression scheme is:

$$B_s = R_s \times Weight_{R_2} + F_s \times Weight_{F_2}$$
$$0 < Weight_{R_2} + Weight_{F_2} < 1 \qquad (3)$$

where the weight is determined in two cases: If $R_s < F_s$, assign the value of $Weight_{R_2}$ to 1, and then $Weight_{F_2} = R_s - 1$, thus $B_s = R_s + F_s \times (R_s - 1)$; If $R_s \geqslant F_s$, assign the value of $Weight_{F_2}$ to 1, and then $Weight_{R_2} = F_s - 1$, thus $B_s = R_s(F_s - 1) + F_s$.

In addition, the position of the fusion prediction bounding boxes should be corrected.

First, normalize the scores of the two output bounding boxes proportionally according to the following scheme:
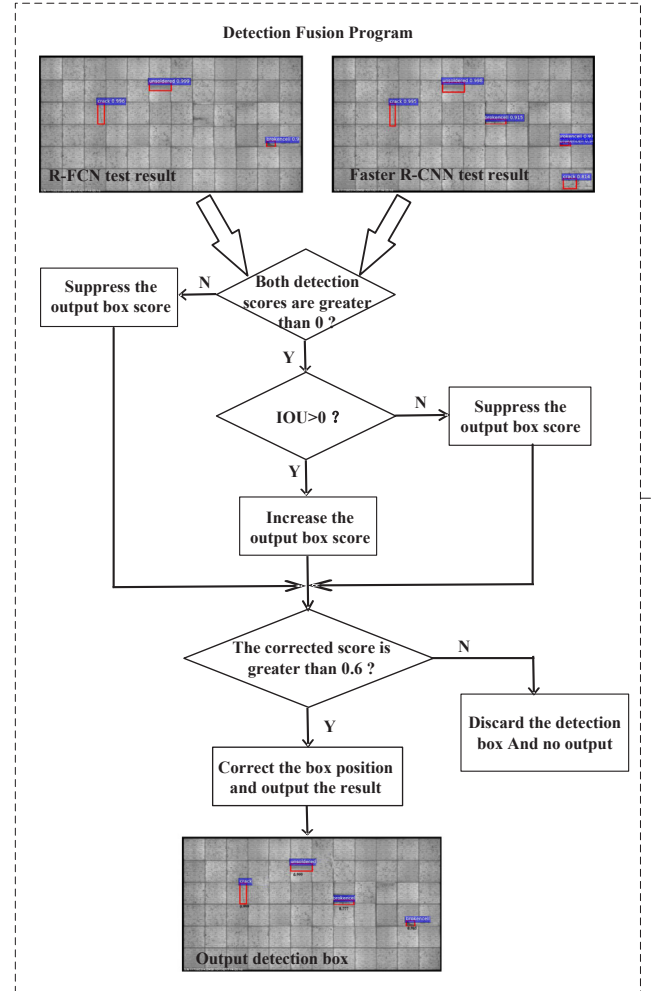
Fig. 8. Schematic diagram of fusion scheme.

$$\begin{cases} s_1 = R_s/(R_s + F_s) \\ s_2 = F_s/(R_s + F_s) \end{cases} \qquad (4)$$

The normalized ratio is used as the weight of the center point position, width and height of the fusion prediction bounding boxes to output more accurate position.

5

Second, connect the central points of the two output prediction bounding boxes. Let

$$
\begin{cases}
R_x = R_b(x) + R_b(w)/2 \\
R_y = R_b(y) + R_b(h)/2 \\
F_x = F_b(x) + F_b(w)/2 \\
F_y = F_b(y) + F_b(h)/2
\end{cases}
\tag{5}
$$

Where x and y represent the horizontal and vertical coordinates, respectively, of the starting point of the upper left corner of the prediction bounding box; w and h represent the width and height, respectively, of the prediction bounding box. (R_x, R_y) represent the center point position of the prediction bounding box, and (F_x, F_y) represent the center point position of the prediction bounding box $F_b(x, y, w, h)$.

Third, calculate the position coordinates of the prediction bounding box after fusion using the coordinates of the prediction bounding box detected by R-FCN and Faster R-CNN, and the normalized prediction bounding box scores are used as the weights. The calculation process is as follows:

$$
\begin{cases}
B_b(x) = (F_x - R_x) * s_2 + R_x \\
B_b(y) = (R_y - F_y) * s_1 + F_y \\
B_b(w) = R_b(w) * s_1 + F_b(w) * s_2 \\
B_b(h) = R_b(h) * s_1 + F_b(h) * s_2
\end{cases}
\tag{6}
$$

4. Results and discussion

Since no research team has shared the solar cell image dataset to date, in order to verify the effectiveness of the proposed algorithm, we constructed a dataset that contained 1462 solar cell surface images in accordance with the method of reference [20–22]. These images are directly obtained by EL imaging system without manually added noise. The dataset contained both defective and defect-free images, and the resolution of the images was uniformly 5232 × 2720. Among them, there are 1461 defective images and 1 defect-free image. Among the 1461 defective images, 281 images only had broken cell defects, 469 images only had crack defects, and 653 images only had unsoldered-area defects. There were 58 mixed defect images, of which 51 had broken cell and crack defects, 3 images had broken cell defects and unsoldered area defects, and 4 contained crack type and unsoldered area defects. Sixty percent of the dataset was selected randomly as the training dataset (consisting of 861 images), and the other forty percent was used as the test dataset (consisting of 600 images). In order to accurately evaluate the detection accuracy of the algorithm, it was necessary to calculate the coincidence between the detection results and the real defects. In this paper, we referred to the labeling format of the PASCAL_VOC dataset in reference [20], and we manually labelled each image in the training set as Ground Truth (GT). Table 3 shows the specific composition of the dataset. In this experiment, the dataset was trained by Faster R-CNN, R-FCN, and YOLO based networks [23–25]. The test environment was a 64-bit Linux system, a Quadro M4000 graphics card (8 g memory), GPU (GTX1080) acceleration and Python Language programming, C++ and MATLAB nesting in the bottom layer, and the Caffe platform to realize the connection and operation before each layer of the network

4.1. Defect detection results

Fig. 9 shows the different defect detection results using Faster R-CNN, R-FCN, and the proposed algorithm in this paper. The three methods can detect three types of defects, i.e., broken cells, cracks, and unsoldered areas. (The different defects that can be detected are labeled in Fig. 9.) Specifically, Fig. 9(a) shows the detection results of Faster R-CNN. ① is a crack defect detected. Obviously, the crack defect is slender, and the position marked by the bounding box is not accurate enough, and the bounding box was too large compared to the size of the defect.

Table 3
Dataset composition.

Defect type	Number of defects in one image	Number of defect images	Total number of defect images
Broken cell	1	319	335
	2	14	
	3	2	
Crack	1	442	524
	2	60	
	3	12	
	4	4	
	5	1	
	6	2	
	7	1	
	13	1	
	14	1	
Unsoldered area	1	559	660
	2	84	
	3	16	
	5	1	

④ is a broken cell that was detected, and the problem of giving repeated bounding box appears; ⑤ is a crack defect that was detected. The gray parts of the defect area and the background area are very close, which caused the detection result to be a false positive area. Fig. 9(b) shows the detection results of R-FCN. The area indicated by the dotted circle box is a crack defect. The defects in this area are relatively hidden, and the features are not obvious. The defect area is very similar to the background area, which leads to the detection result being a false negative area. Fig. 9(c) shows the detection results of this proposed method. It can be seen that cracks, unsoldered areas, and broken cells are detected accurately and with high confidence at ①, ②, and ④, respectively. ③ is a broken cell defect that was detected with relatively low confidence. Generally speaking, the detection results of this method are superior to the other two algorithms due to its high detection accuracy, low false positive rate, low false negative rate, and high confidence value.

Fig. 10 shows a partial enlarged view of the three defect detection results obtained by this method. Fig. 10(a) shows the detected broken cells. Fig. 10(b) shows the detected cracks. Fig. 10(c) shows the detected unsoldered areas. It can be seen that the proposed method can accurately detect all three types of defects. Since we set a new size and ratio to the candidate box, the accuracy of the position of the bounding box was relatively high. The advantages of combining Faster R-CNN and R-FCN also make the detection results of the proposed method have a higher confidence value.

In order to verify the effectiveness of the proposed method further, we also compared and analyzed the regression accuracy of the three defect bounding boxes. Fig. 11 shows the partial enlarged image of the bounding box for broken cells of different defect models. Comparing the bounding box for broken cells obtained by Faster R-CNN, R-FCN, and the proposed method with the ground truth bounding box, it was apparent that, although Faster R-CNN can detect broken cells, the regression accuracy error of the bounding box was relatively large, and the IOU value between the bounding box of broken cell defect and the real box was low. Due to the combination of the advantages of high detection accuracy of R-FCN, the method proposed in this paper had a better performance. It can accurately detect the broken cells and also obtain a bounding box with high regression accuracy.
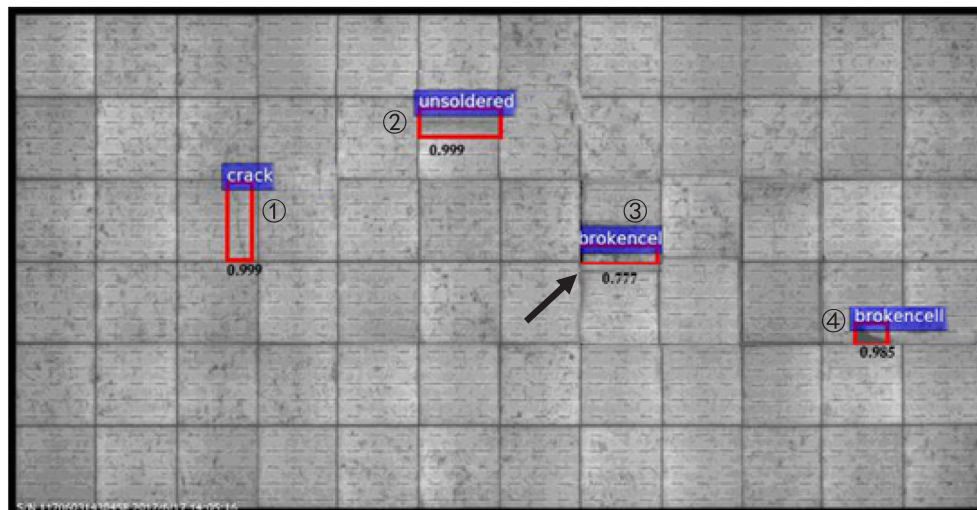
Fig. 12 shows the partial enlarged image of the bounding box for cracks of different defect models. It is apparent that, although Faster R-CNN accurately outlines the location of cracks, the bounding box is relatively large and the regression error is large. Both R-FCN and the method proposed in this paper were very effective because they were able to detect cracks and also obtain the bounding box with high-accuracy regression. Fig. 13 shows the partially-enlarged image of the bounding box for unsoldered areas of different defect models. It is
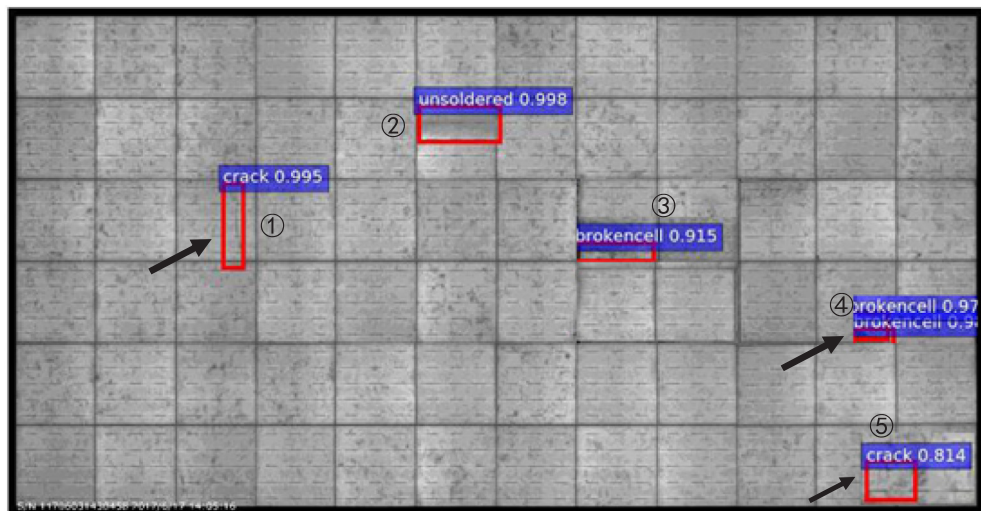
（a）The defect detection results of Faster R-CNN
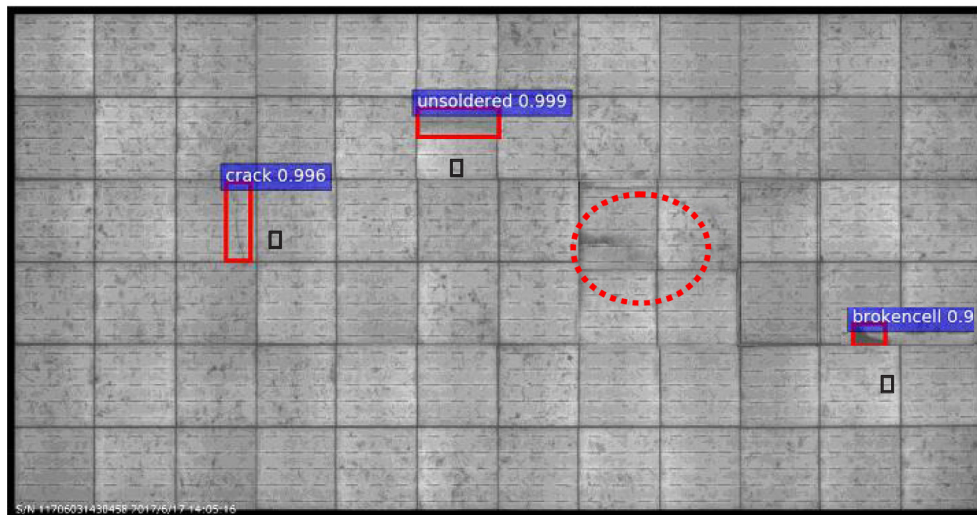
（b）The defect detection results of R-FCN

（c）The defect detection results of the proposed method

Fig. 9. The different defect detection results using Faster R-CNN, R-FCN and the proposed algorithm.

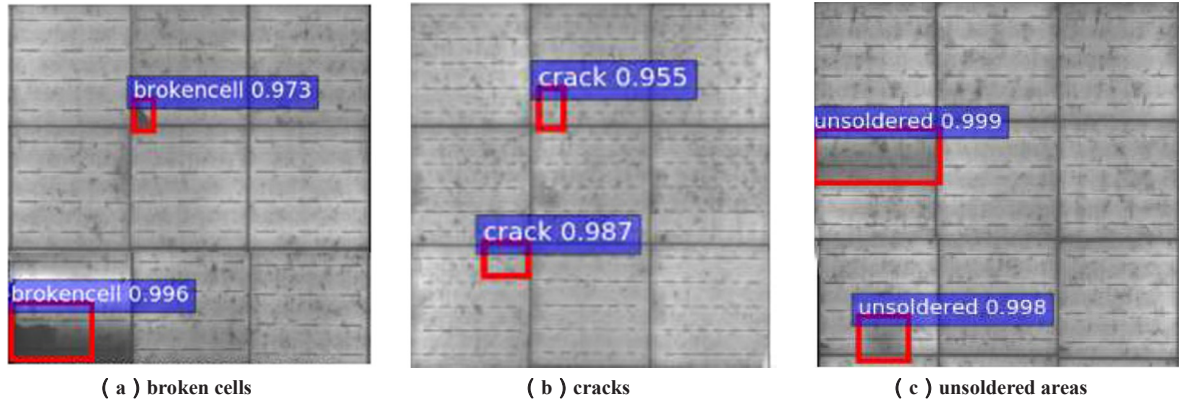(a) broken cells　　　　　　　　(b) cracks　　　　　　　　(c) unsoldered areas

Fig. 10. The partial enlarged images of the three defect detection results obtained by this method.

apparent that, compared with the other two methods, the detected bounding box of Faster R-CNN was either too large or too small, and the accuracy of the regression was not high. The new method presented in this paper had higher regression accuracy.

4.2. Quantitative analysis

This section presents a quantitative analysis of the detection performance of Faster R-CNN, R-FCN and the method of this paper for three types of defects using three indicators, i.e., false negative rate, false positive rate, and detection accuracy. Table 4 shows the statistics of the defect detection results of the three methods. Six hundred images of the test set were used in the three methods, among which 200 images were used as the test set to detect each defect. In general, Faster RCNN had more false positive images than false negative images, but the number of false negative images was greater than the number of false positive images for R-FCN. The new method in this paper combines the advantages of the two, and the numbers of false positive images and false negative images were reduced significantly. Specifically, for Faster R-CNN, the number of false negative images of the three defects was higher than the other two methods. For R-FCN, the number of false negative images of the three defects was higher than the other two

methods. For different defects, the detection effects were different, which was related to the characteristics of the cracks. The gray-level change of the cracks was not obvious in the EL image of the solar cell, and it is difficult for the human eye to recognize some cracks. Because the proposed method integrated the advantages of Faster R-CNN and R-FCN and modified the candidate box, the total number of false positive images and false negative images of our new method was reduced significantly. The number of false negative images for unsoldered areas was 0, and the number of false positive images for unsoldered areas was as low as 3. The remaining two defects had false negative and false positive images, but the numbers were very small.

Table 5 shows the statistical results of the detection accuracies, false positive rates, and false negative rates of the three methods. It is apparent that the detection accuracy of Faster R-CNN for broken cells was high, but the detection accuracy for cracks was low, and the false negative rate for unsoldered areas was higher. Overall, the false negative rate of Faster R-CNN was greater than the false positive rate. The detection accuracy for cracks and unsoldered areas of R-FCN was higher than that of Faster RCNN. However, for broken cells, the detection accuracy was lower than that of Faster RCNN. In general, R-FCN had a higher false positive rate than false negative rate. The detection accuracy of our proposed method was higher than that of R-FCN, and the
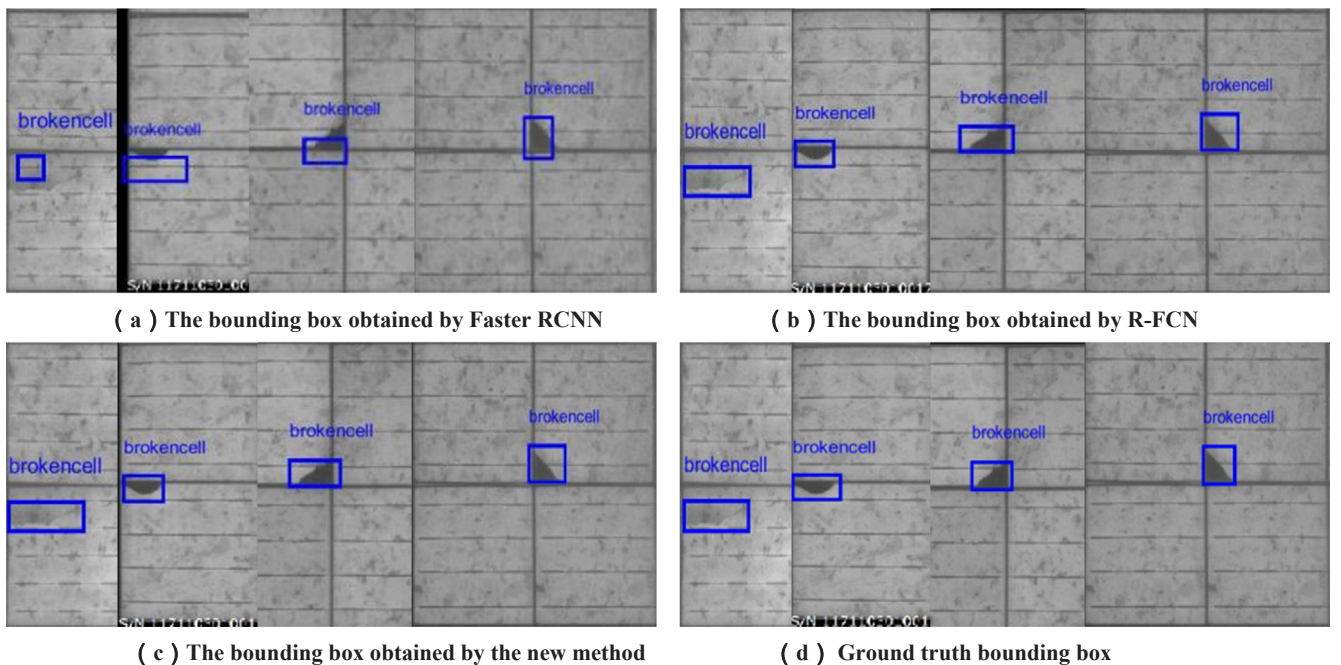

(a) The bounding box obtained by Faster RCNN

(b) The bounding box obtained by R-FCN

(c) The bounding box obtained by the new method

(d) Ground truth bounding box

Fig. 11. The partial enlarged image of the bounding box for broken cells of different defect models.

(ａ) The bounding box obtained by Faster RCNN

(ｂ) The bounding box obtained by R-FCN

(ｃ) The bounding box obtained by the new method
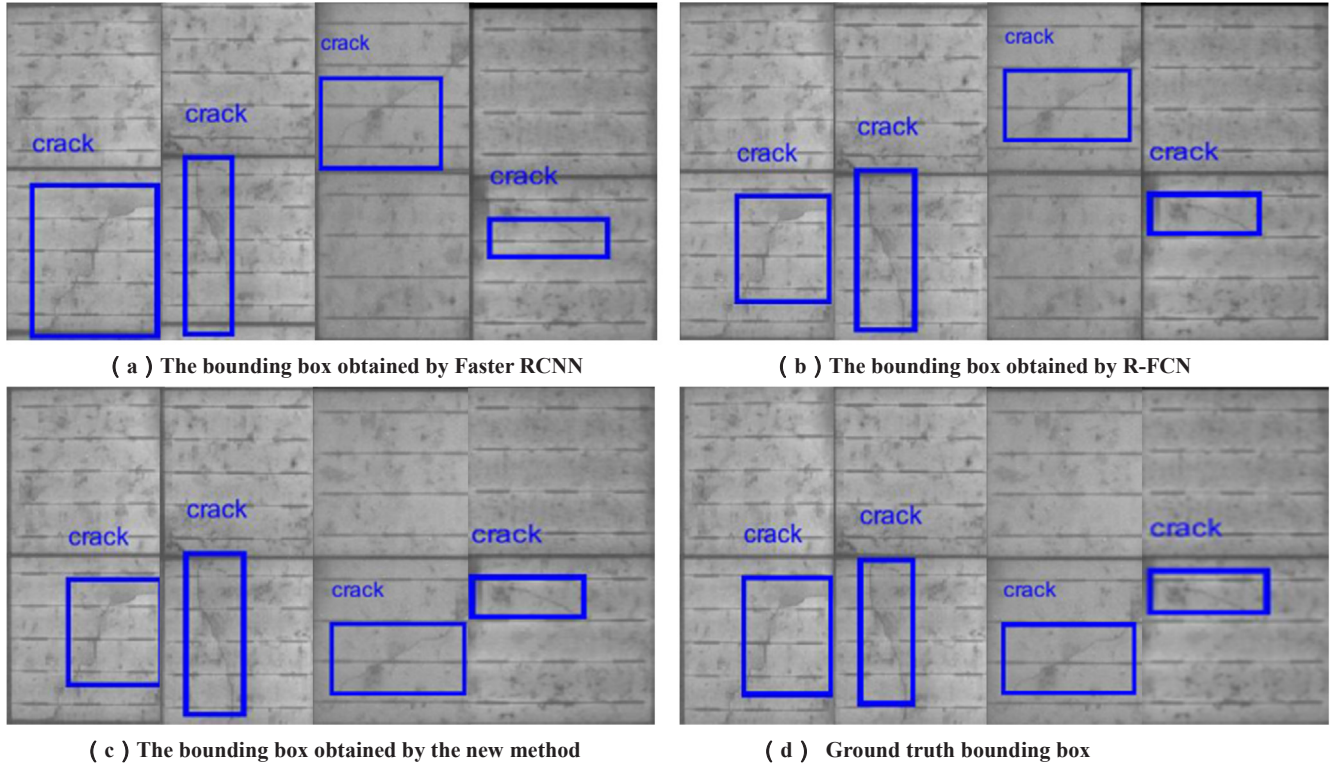
(ｄ) Ground truth bounding box

Fig. 12. The partial enlarged image of the bounding box for cracks of different defect models.

false negative and false positive rates were decreased significantly.

In order to intuitively analyze the detection accuracy of the three methods for different defects, Fig. 14 shows a bar chart of detection accuracy of three defects using different detection models. It is apparent that the three methods have high detection accuracy for broken cells, but the detection accuracies of Faster R-CNN for cracks and unsoldered

areas were low. Due to the use of the candidate box more suitable for solar cells, the detection accuracy of the proposed method was improved to some extent, especially for the defects of unsoldered areas. Compared with Faster R-CNN, the detection accuracy of the proposed method was improved by 7%. Compared with R-FCN, the detection accuracy of the proposed method was improved by at least 3%. The
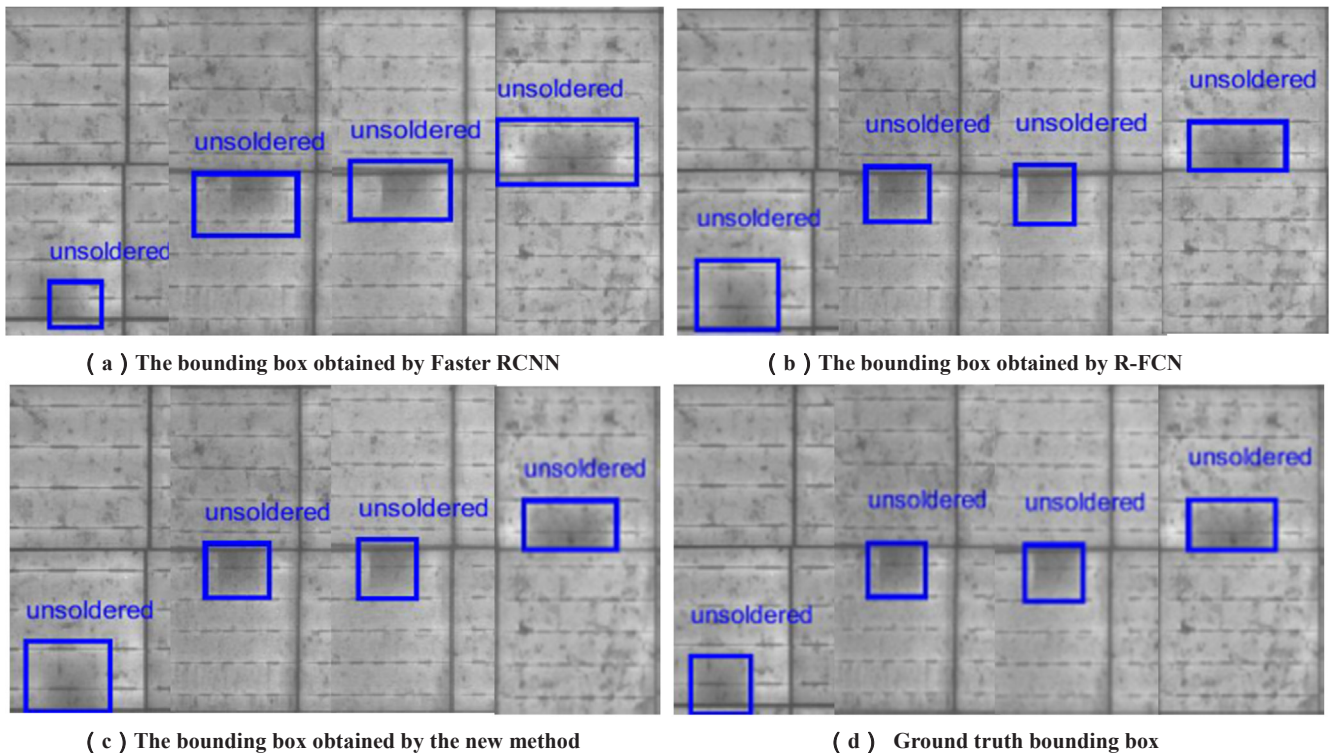

(ａ) The bounding box obtained by Faster RCNN

(ｂ) The bounding box obtained by R-FCN

(ｃ) The bounding box obtained by the new method

(ｄ) Ground truth bounding box

Fig.13. The partial enlarged image of the bounding box for unsoldered areas of different defect models.

Table 4
The defect detection results statistics of the three methods.

	Faster R-CNN			R-FCN			The proposed method		
	Number of accurate detection images	Number of false positive images	Number of false negative images	Number of accurate detection images	Number of false positive images	Number of false negative images	Number of accurate detection images	Number of false positive images	Number of false negative images
Broken cell	190	1	9	188	8	4	198	1	1
Crack	179	5	16	192	6	2	195	3	2
Unsoldered area	179	0	21	188	5	7	197	0	3
Total number	548	6	46	568	19	13	590	4	6

proposed method had better detection performance for detecting specific small-area objects.

Table 6 provides the statistics for the regression accuracy of the bounding box for three defects using different methods. Faster-RCNN had a high regression accuracy for the bounding box for unsoldered areas. However, compared with the other two methods, the regression accuracy of bounding box was poor. The regression accuracies of the bounding box for broken cells and cracks were less than 50%. R-FCN had high regression accuracy of bounding box for broken cells and unsoldered areas. Even though R-FCN's regression accuracy of the bounding box for cracks of R-FCN was low, it was higher than that of Faster-RCNN. The proposed method performed well and had a higher regression accuracy than that of R-FCN for the bounding box with three types of defects. In order to more intuitively display the regression accuracy of the bounding box for the three types of defects for the different algorithms, Fig. 15 shows a Bar chart of the regression accuracy of bounding box for the three types of defects for the different detection models.
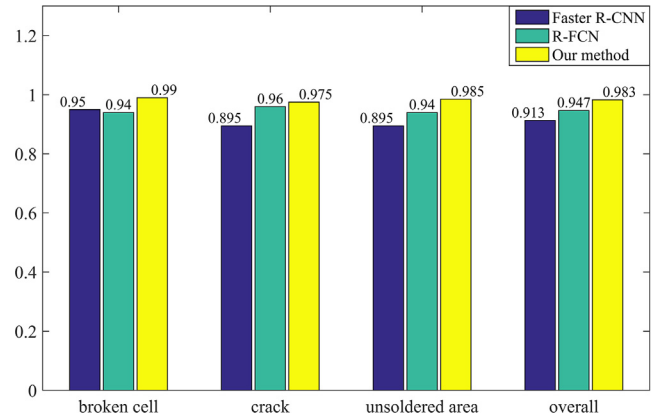
To determining the optimal value of the parameters ζ in formula (1) using experimental methods, we selected 180 defective images (including 60 images with broken cells, 60 images with cracks, and 60 images with unsoldered areas) to test. Fig. 16 shows the change curve of false positive rate and false negative rate with the increase of parameter ζ. The red line indicates the change of the false positive rate, which decreases with the increase of parameter ζ. The blue line indicates the false negative rate, which increases with the increase of parameter ζ. The green line in the figure shows the change of the average error rate. It can be seen that the average error rate is the lowest when the parameter is 0.9. Therefore, the parameter ζ is set to 0.9 in this paper

In addition, in this paper, we compared the detection effect of different methods on the surface defects of the same single solar cell EL image. Images in the test set were detected using YOLO (You Only Look Once) [23], YOLO V9000 [24], YOLO V3 [25], Faster R-CNN, and R-FCN, as well as the method proposed in this paper, and the detection results are shown in Table 7. YOLO turns the whole detection problem into a regression problem. The network structure is simple, and the detection speed is high, but the accuracy of detection is low compared with other similar methods in the table. YOLO V9000 and YOLO V3 are improved versions of the YOLO model. The detection accuracy is sorted from small to large in order as YOLO < YOLO V900 < YOLO V3. R-FCN uses the Faster R-CNN architecture, but it only contains the convolution network, which reduces the computation and improves the

Fig. 14. Bar chart of detection accuracy of three defects using different detection models.

Table 6
Statistics on the regression accuracy of bounding box for three defects using different method.

Defect type	Faster-RCNN	R-FCN	The proposed method
Broken cell	0.4693	0.7519	0.7785
Crack	0.4770	0.7372	0.7535
Unsoldered area	0.6950	0.8110	0.8337
Total	0.5471	0.7376	0.7886

detection rate compared with Faster R-CNN. As can be seen from the table, the detection accuracy of this method is improved greatly compared with other methods under the condition that real-time detection basically is achieved. Fig. 17 shows the class-agnostic precision-recall curves, and the figure shows that, compared with YOLO, YOLO v9000, YOLO v3, Faster R-CNN, and R-FCN, the proposed method can improve the detection accuracy and achieve better regression accuracy of the bounding box results.

The detection results of Faster R-CNN on the test set using different settings of anchors. The network is VGG-16, and the training data are trainval. The default setting of using 5 scales and 5 aspect ratios (73.5%) is the same as that in Table 8. Each image in the training set contained at least one defect, and some images contain multiple defects,
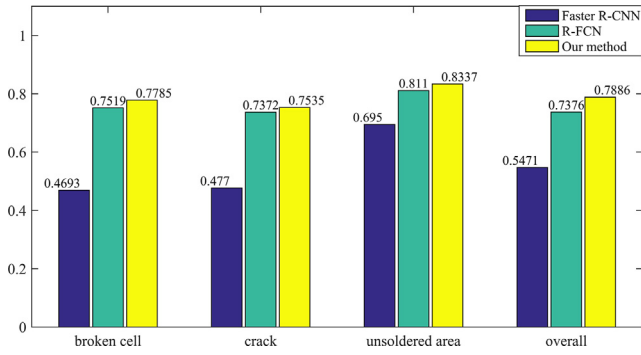
Table 5
The statistical results of the detection accuracy, false positive rate and false negative rate of the three methods.

	Faster R-CNN(%)			R-FCN(%)			The proposed method(%)		
	Detection accurate rate	False positive rate	False negative rate	Detection accurate rate	False positive rate	False negative rate	Detection accurate rate	False positive rate	False negative rate
Broken cell	95	0.5	4.5	94	4	2	99	0.5	0.5
Crack	89.5	2.5	8	96	3	1	97.5	1.5	1
Unsoldered area	89.5	0	10.5	94	2.5	3.5	98.5	0	1.5
Total number	91.3	1	7.7	94.7	3.2	2.1	98.3	0.7	1

Fig. 15. Bar chart of the regression accuracy of bounding box for three defects using different detection models.
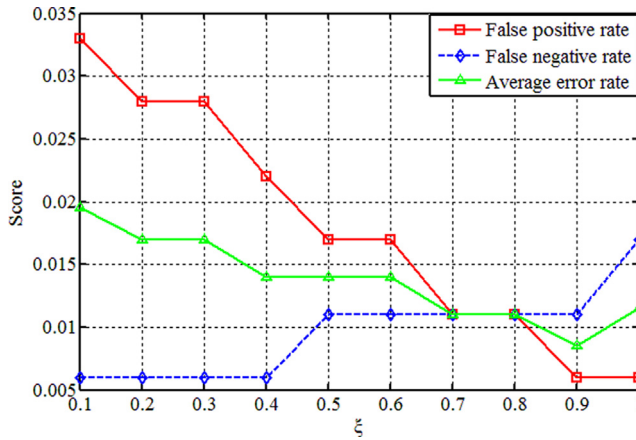

Fig. 16. The change curves of false positive rate and false negative rate with the increase of parameter ζ.

Table 7
The map value statistics for different methods.

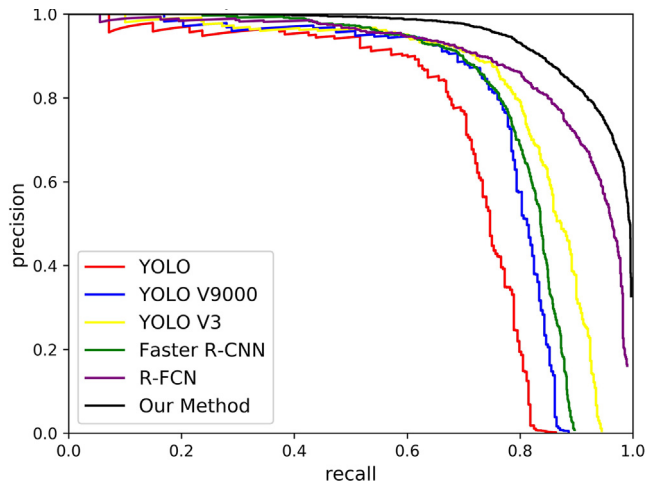| Method	Map(%)
Yolo	74.2
Yolo V9000	79.7
Yolo V3	82.5
Faster R-CNN	80.1
R-FCN	82.9
Our Method	85.7

Fig. 17. Class-agnostic precision-recall curves.

Table 8
The anchor parameters statistics.

Setting	Anchor scales	Aspect rations	Map(%)
1 scale 1 ration	256^2	1:1	68.8
3 scale 3 ration	$[128^2, 256^2, 512^2]$	[1:1, 1:2, 2:1]	72.4
5 scale 5 ration	$[32^2, 64^2, 128^2, 256^2, 512^2]$	[1:1, 1:2, 1:3, 3:1, 2:1]	76.5

i.e., each image contains different numbers of real boxes. All anchors of each image in the training set were examined. Table 8 shows the statistics concerning the anchor parameters. For broken cells, cracks, and unsoldered areas, utilizing the anchor of 2562, 1:1 will make the area of each anchor much larger than the area of the defects, which causes a small percentage of defects in the anchor. However, the surface defects of solar cells have unique slender and small characteristics, which make the anchor and GT overlap slightly during the training process. In turn, the anchor was judged as a negative sample or a non-negative and non-positive sample, and it was discarded. Based on the size of the defects in the training set and the ratio of the length and width of the GT bounding box, this article improves the scale and ratios of the anchor. There were 25 reselected anchors (ratios*scale). There are 5 types of ratios and 5 types of scales. See Table 8 for specific parameters.

In order to further verify the effectiveness of different strategies in the proposed method, in this paper, we used different strategies to train and test the network. Table 9 shows the strategy that was used and the detection results after training. Table 9 shows that the design decisions listed in this article will result in a significant increase in the value of mean Average Precision (mAP). Adding Online Hard Examples Mining can increase the Map value by nearly 3%. Multi-Channel Conv-fusion was added to increase the Map value by 6.5%. This proves the effectiveness of the method proposed in this paper. In summary, the proposed method is an effective method for detecting defects on the surfaces of solar cell EL images.

Finally, in order to evaluate the real-time index of the proposed method, we use Faster RCNN, Faster RCNN method with improved anchor, RFCN and this method to test 600 solar cell images with defects. Table 10 records the total time required for testing. Because the method proposed in this paper adopts multi-channel joint detection, the detection time required is more than that of other methods, indicating that the real-time performance of the method in this paper is relatively poor, but considering that the average detection time per image is less than 1 s, its detection time can still meet the needs of real-time applications.

5. Conclusions

According to the characteristics of the defects on the surfaces of solar cells, we proposed a solar cell surface defect detection method based on the fusion model of Faster R-CNN and R-FCN. The accuracy of the detection was improved by the complementary fusion of the detection results of the two models. The proposed method also effectively overcame the problems that 1) the accuracy of defect's position regression for Faster R-CNN is low due to the number limitation of the anchor points. 2) The positive and negative samples are unbalanced,

Table 9
The the detection effect statistics Using different strategies to train network.

Setting			
Add multi-scale and multi-ratios	√	√	√
Online Hard Examples Mining		√	√
Multi-Channel Conv-fusion			√
Map (%)	76.5	79.2	85.7

Table 10
The total testing time statistics.

Methods	Testing time (s)
Faster RCNN	254.146
Faster RCNN (with improved anchor)	270.817
RFCN	231.878
Our method	382.793

i.e., the negative sample space is much larger than the defect area in solar cells image. Thus, the Multi-scales Multi-ratios strategy and the hard negative sample mining strategy were adopted to enhance the discrimination ability of the algorithm and to improve the accuracy of detection. The experimental results showed that the fusion scheme proposed in this paper effectively overcomes the problem of high false negative rate and false positive rates of a single network to detect defects, and it greatly improves the detection accuracy of defect locations while improving the recall rate of the object. The next research focus is to estimate the geometric size of the defects on this basis.

Acknowledgement

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.infrared.2020.103334.

References

[1] N.S. Lewis, Research opportunities to advance solar energy utilization, Science 351 (6271) (2016) aad1920.

[2] D.M. Tsai, S.C. Wu, W.C. Li, Defect detection of solar cells in electroluminescence images using Fourier image reconstruction, Sol. Energy Mater. Sol. Cells 99 (2012) 250–262.

[3] D.M. Tsai, J.Y. Luo, Mean shift-based defect detection in multicrystalline solar wafer surfaces, IEEE Trans. Ind. Inf. 7 (1) (2010) 125–135.

[4] Y.C. Chiou, J.Z. Liu, Y.T. Liang, Micro crack detection of multi-crystalline silicon solar wafer using machine vision techniques, Sensor Rev. 31 (2) (2011) 154–165.

[5] D.M. Tsai, C.C. Chang, S.M. Chao, Micro-crack inspection in heterogeneously textured solar wafers using anisotropic diffusion, Image Vis. Comput. 28 (3) (2010) 491–501.

[6] X. Qian, H. Zhang, C. Yang, et al., Micro-cracks detection of multicrystalline solar cell surface based on self-learning features and low-rank matrix recovery, Sensor Rev. 38 (3) (2018) 360–368.

[7] D.M. Tsai, G.N. Li, W.C. Li, et al., Defect detection in multi-crystal solar cells using clustering with uniformity measures, Adv. Eng. Inf. 29 (3) (2015) 419–430.

[8] Jian Guan, et al., Adaptive fractional Fourier transform-based detection algorithm for moving target in heavy sea clutter, IET Radar Sonar Navig. 6 (5) (2012) 389–401.

[9] W.C. Li, D.M. Tsai, Wavelet-based defect detection in solar wafer images with inhomogeneous texture, Pattern Recogn. 45 (2) (2012) 742–756.

[10] F.C. Chen, M.R. Jahanshahi, NB-CNN: deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion, IEEE Trans. Ind. Electron. 65 (5) (2017) 4392–4400.

[11] T.H. Sun, F.C. Tien, F.C. Tien, et al., Automated thermal fuse inspection using machine vision and artificial neural networks, J. Intell. Manuf. 27 (3) (2016) 639–651.

[12] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, Science 313 (5786) (2006) 504–507.

[13] R. Girshick, J. Donahue, T. Darrell, et al., Rich feature hierarchies for accurate object detection and semantic segmentation, in: IEEE Conference on Computer Vision and Pattern Recognition, IEEE Comput. Soc., 2014, pp. 580–587.

[14] R. Girshick, Fast r-cnn, in: Proceedings of the IEEE international conference on computer vision, 2015, pp. 1440–1448.

[15] S. Ren, K. He, R. Girshick, et al., Faster r-cnn: Towards real-time object detection with region proposal networks, Adv. Neural Inf. Process. Syst. (2015) 91–99.

[16] J. Dai, Y. Li, K. He, et al., R-fcn: Object detection via region-based fully convolutional networks, Adv. Neural Inf. Process. Syst. (2016) 379–387.

[17] H. Chen, Y. Pang, Q. Hu, et al., Solar cell surface defect inspection based on multispectral convolutional neural network, J. Intell. Manuf. (2018) 1–16.

[18] A. Bartler, L. Mauch, B. Yang, et al., Automated Detection of Solar Cell Defects with Deep Learning, in: 2018 26th European Signal Processing Conference (EUSIPCO), IEEE, 2018, pp. 2035–2039.

[19] K. He, X. Zhang, S. Ren, et al., Spatial pyramid pooling in deep convolutional networks for visual recognition, IEEE Trans. Pattern Anal. Mach. Intell. 37 (9) (2015) 1904–1916.

[20] M. Everingham, The PASCAL visual object classes challenge 2007, 2009, http://www. pascal-network. org/challenges/VOC/voc2007/workshop/index.html.

[21] M. Everingham, L. Van Gool, C.K.I. Williams, et al., The pascal visual object classes (voc) challenge, Int. J. Comput. Vision 88 (2) (2010) 303–338.

[22] X. Chen, A. Gupta, An implementation of faster rcnn with study for region sampling, 2017. arXiv preprint arXiv:1702.02138.

[23] J. Redmon, S. Divvala, R. Girshick, et al. 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.

[24] J. Redmon, A. Farhadi, YOLO9000: better, faster, stronger, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7263–7271.

[25] J. Redmon, A. Farhadi, Yolov3: An incremental improvement, 2018. arXiv preprint arXiv:1804.02767.