Surface defect detection of solar cell based on similarity non-maximum suppression mechanism

Yanling Wang (✉ ylwang1981@126.com)

 School of information, Shanxi University of Finance and Economics

Ting Hou

 School of Electronic Information Engineering, Taiyuan University of Science and Technology

Xiong Zhang

 School of Electronic Information Engineering, Taiyuan University of Science and Technology

Hong Shangguan

 School of Electronic Information Engineering, Taiyuan University of Science and Technology

Pengcheng Zhang

 Shanxi Provincial Key Laboratory for Biomedical Imaging and Big Data, North University of China

Jie Li

 School of information, Shanxi University of Finance and Economics

Baoyue Wei

 School of Electronic Information Engineering, Taiyuan University of Science and Technology

Abstract

The surface defects such as cracks, broken cells and unsoldered areas on the solar cell caused by manufacturing process defects or artificial operation seriously affect the efficiency of solar cell. For the surface defects of solar cell, which have the characteristics of various shapes, large scale changes, and difficult to detect, a surface defect detection algorithm based on similarity non-maximum suppression mechanism is proposed by improving the Faster region-based convolution neural network (R-CNN) in this paper. In the proposed algorithm, a similarity non-maximum suppression mechanism is used, and the effectiveness of prediction frame screening is improved by introducing the cosine similarity of candidate box aspect-ratio. In addition, the cross-layer connection based on Shuffle operation and the three-branch dilated convolution block are introduced in the main feature extraction channel, which improves the network's ability to express features through multi-scale feature fusion. The experimental results show that, compared with the latest deep learning target detection models, the proposed algorithm not only has higher detection accuracy, but also lower false detection and missed detection rates in various types of defect detection.

1 Introduction

The widespread application of solar photovoltaic technology has well alleviated the environmental problems caused by the excessive use and consumption of non-renewable energy sources such as coal and oil. However, the production process of solar panels is complex and the substrate is fragile, and process defects or human errors can easily lead to subtle, hidden and hard-to-detect defects such as cracks, broken cells and unsoldered areas on the surface of solar cells [1]. The existence of solar cell surface defects can seriously reduce the photoelectric conversion efficiency and service life of the cells. The in-depth research of solar cell surface defect detection technology and timely replacement of damaged panels based on the detection results are the keys to improving the efficiency of solar photovoltaic power generation [2].

Traditional solar cell surface defect detection methods contain laser scanning method [3], acoustic wave method [4] and Hertzian spectroscopy method [5], but the preprocessing is time-consuming and labor-intensive, and the detection accuracy is low. The defect detection methods based on machine vision have achieved rapid development due to their advantages of convenient operation, high detection accuracy and good real-time performance. The application of clustering methods [6], matrix decomposition methods [7], and wavelet transform methods [8] has greatly improved the detection accuracy of solar cell surface defects, but the traditional image domain and transform domain detection methods need to manually obtain relevant target feature information from the original input, and there are still many limitations in portability, accuracy and real-time performance.

Since Hinton et al. proposed the use of neural networks to automatically learn high-level features in multimedia data [9], the application of deep learning in the field of image object detection has achieved great success [10]. In 2014, Girshick et al. firstly used the R-CNN for object detection, which greatly improved the accuracy of object detection [11]. Through similar end-to-end training, not only can deep abstract features be extracted, but also feature extraction, selection, and classification can be completed in the same model, improving both performance and efficiency. According to the presence or absence of the candidate box generation stage, the target detection methods based on deep learning can be divided into two categories: one is a single-stage target detection algorithm that directly calculates the image to generate the detection result, such as the YOLO (you only look once) series [12–16], the detection speed is fast, but the accuracy is low. The SSD (Single Shot Multibox Detector) largely balances the detection speed and accuracy of the single-stage detection algorithm [17]; the other is the two-stage target detection algorithm, which is based on a pre-generated candidate area to do a secondary correction to obtain the detection result, the detection accuracy is high, and the detection speed is slow. R-CNN has laid the foundation of the two-stage target detection algorithm based on deep learning. The algorithm design is ingenious, but the efficiency is low. Based on R-CNN, He et al. and Girshick et al. proposed the spatial pyramid pooling net (SPPNet) [18] and Fast R-CNN algorithm [19] in 2015, respectively. The detection speed was improved because these two methods only need to send the image into the depth network once. The Faster R-CNN proposed by Ren et al. added a region proposal network (RPN) on the basis of Fast R-CNN, which improved the network computing speed [20]. Dai et al. proposed a region based fully convolutional network (R-FCN), which improved detection accuracy through the use of location-sensitive score maps [21]. The Cascade R-CNN algorithm proposed by Cai et al. improved the Faster R-CNN algorithm by cascading multiple detection heads [22]. In order to identify targets of different scales, Girshick et al. proposed a feature pyramid network (FPN), which constructed a feature pyramid through the multi-scale pyramid inherent in a deep convolutional network, and realized the construction of high-level semantic feature maps at all hierarchical scales [23].

With the successful application of deep learning in the field of image target detection, more and more deep learning methods are applied to the surface defect detection of solar cell [24]. Chen et al. designed a visual defect detection method based on multispectral deep convolution neural network (CNN) [25], Zhang et al. made a complementary fusion of Faster R-CNN and R-FCN model, and proposed a solar cell surface defect detection algorithm fused with multi-channel CNN [26]. In view of the characteristics of solar cell surface defects with diverse shapes, large scale changes, subtle features, similar to the background, and difficult detection, this paper takes Faster R-CNN as the baseline model, and proposes a solar cell surface defect detection algorithm based on similarity non-maximum suppression mechanism. The proposed algorithm has improved the network structure according to the characteristics of CNN, so that it can use different sizes of receptive fields for targets of different scales while detecting low-level and high-level features, thus effectively improving the network performance and test results. The specific contributions are summarized as follows:

(1) The use of cross-layer connections in Resnet [27] effectively improves the utilization of features, but downsampling and pooling operations easily lead to damage to image features and increase the amount of network parameters. In this paper, a cross-layer connection based multi-feature fusion network (CCMFN) is proposed. In the feature extraction stage, a cross-layer connection based on shuffle operation is introduced. Adding shuffle operation before feature map fusion can adjust the size of the image without destroying the features of the input image and without increasing the complexity of the operation.

(2) In the detection methods based on Faster R-CNN, the features output from a fixed convolutional layer are often used to characterize the defects of different morphologies and scales, and the detection effect is not good. The proposed algorithm improves the RPN network used to generate candidate boxes in Faster

R-CNN, adding three-branch dilated convolution (TDC) block. By adjusting the dilation factor, the convolution kernel can obtain different receptive fields, the defect features of different scales can be better learned, the more accurate and applicable candidate boxes can be extracted by RPN.

(3) The candidate box screening strategy of Faster R-CNN is the non-maximum suppression (NMS) mechanism, which uses the intersection over union (IOU) between two candidate boxes to evaluate their relationship. This evaluation criterion only considers the overlapping area between two candidate boxes, and cannot fully describe the overlapping relationship of bounding boxes. In order to make the candidate box screening strategy more suitable for solar cell defect detection, this paper proposes a similarity non-maximum suppression (S-NMS) mechanism to improve the effectiveness of prediction box screening by introducing the cosine similarity of candidate box aspect-ratio.

2 Faster R-cnn

As a target detection scheme based on CNN, Faster R-CNN mainly solves two problems: first, the RPN is used to quickly generate candidate regions; second, the RPN and Fast-RCNN networks share parameters through alternate training. The original network structure of Faster R-CNN is shown in Fig. 1. For the input defect detection image, firstly, feature extraction is performed through VGG, Resnet and other networks to obtain the basic feature map; secondly, the RPN network is used to generate many regional candidate boxes, and then the candidate boxes are mapped to the last layer of convolutional feature map; finally, the fully connected layer is used to classify the target specifically and perform accurate bounding box regression.

Due to the existence of a large number of redundant boxes, it is necessary to use NMS for candidate boxes screening after RPN generates many regional candidate boxes. By removing redundant boxes with high confidence, the candidate boxes with high-quality are retained for bounding box regression and classification. According to the target bounding box list and its corresponding confidence score list, NMS deletes bounding prediction boxes with large overlap by setting a threshold.

3 Method

In order to make the detection network have better feature extraction capability for solar cell surface defects of different morphologies and different scales, this paper takes Faster R-CNN as the baseline model, and proposes a network based on multi-feature fusion and S-NMS. The overall framework of the network is shown in Fig. 2. Firstly, the basic feature map with rich semantic features is obtained by CCMFN. Then, the features of different receptive fields at different scales are fully fused by TDC block, and then the candidate boxes are generated. Finally, more accurate detection boxes are screened out by using S-NMS in the testing phase, which further improves the accuracy of detection. The principles of CCMFN, RPN multi-scale receptive field fusion network, and S-NMS modules are described below, respectively.

3.1 Cross-layer connection multi-feature fusion network (CCMFN)

Aiming at the characteristics of extremely small surface defects of solar cells, high similarity with background, high detection difficulty and low regression accuracy, this paper proposes CCMFN. The network structure diagram and network visualization flow chart are given in Fig. 2(B1) and (B2), respectively. The images to be detected are input into the pre-training model, and the main feature maps containing defective parts are obtained after passing through VGG16 and Resnet101 networks, respectively. In order to effectively preserve the effective features of different layers of the network and avoid the problem of gradient disappearance and degradation caused by the deepening of the network, the idea of identity in Resnet is introduced in this paper. The features of the second and third layers of the network are extracted, converted into features of the same size as the features of the fifth layer by Shuffle operation, and cascaded. Then, the dimension is reduced by 1×1 convolution, and the basic feature map is finally obtained. In the process of adjusting the size of the feature map, the shuffle operation not only retains the information of different scales, but also does not introduce additional parameters, which greatly speeds up the running speed of the network. The design retains the shallow detailed texture information and the deep rich semantic information without increasing the network operation cost, which effectively improves the regression accuracy and detection accuracy of the network.

3.2 RPN multi-scale receptive field fusion network

In Faster R-CNN, after the basic feature map is input to RPN, a 3×3 convolutional layer is first used for further feature fusion, and then the candidate boxes are generated. According to the study in literature [28], the detection performance of the network on targets with different scales is greatly affected by the receptive field, and the size of the target has a strong correlation with the appropriate receptive field. In general, the deeper the network level, the larger the receptive field, and the stronger the ability to handle large targets. The shallower the network level, the smaller the receptive field, and the stronger the ability to handle small targets.

The three types of defects in the solar dataset have subtle morphology and low discriminability, and the scales of defects of the same type are inconsistent. In order to extract features of different scales more comprehensively, this paper designs a RPN feature fusion network composed of TDC blocks. The structure is shown in Fig. 2(C). The single-scale feature images are taken as input, and the Resnet structure shown in Fig. 3 is taken as an example, in a single residual block of the bottleneck layer, a feature map of a specific scale is created through parallel branches, and different dilation rates are used in the 3×3 convolutional layer to form different receptive field branches, the branch with large receptive field extracts large-scale defect features, the branch with small receptive field extracts small-scale defect features, and different branches adopt the same structure and share weights. In addition, the use of zero-padding on the convolution kernel enables the network to achieve an increase in the receptive field without increasing the network computation.

3.3 Similarity non-maximum suppression (S-NMS)

In Faster R-CNN, after the network generates a large number of region candidate boxes through RPN, NMS is used as a post-processing algorithm to remove the redundant boxes so that the network performance can strike a balance between recall rate and precision. The traditional NMS algorithm is based on a pre-set threshold, which deletes the detection results that are larger than the threshold, otherwise keeps them. If the threshold is reduced to a certain level, the false positive rate will increase. However, when the evaluation is performed with a high overlap threshold, the lower scoring detection boxes are completely removed so that the false detection rate increases, which in turn leads to a decrease in the average precision measured over a range of overlapping thresholds. The later improvements are all made for the case of target crowding and occlusion [28, 29], while there is no occlusion in the solar cell surface defects used in this paper, but there is a dense phenomenon of targets in part. In order to further improve the detection accuracy and precision, this paper proposes the S-NMS criterion based on the cosine similarity of candidate box aspect-ratio.

In the testing process, for the set B of each category containing the initial prediction box, the classification execution degree S corresponding to each prediction box is ranked from high to low, and then the highest score m and its corresponding prediction box Bm are deleted from set B and placed in the empty set D. Calculate the IOU value and the cosine similarity sim(X,Y) between the remaining prediction box and Bm in set B in turn, and add the IOU value and the cosine similarity of diagonal sim(X,Y) to get Ω. If Ω exceeds the set threshold Nt, the corresponding prediction box will be deleted, otherwise, it will be put into set D. The above process is repeated until the set B is empty, and the input list is returned.

The S-NMS method improves the screening accuracy of the traditional NMS algorithm by introducing the prediction box similarity as another criterion in the threshold criterion, which in turn improves the detection accuracy and classification accuracy of the network.

4 Experimental Results And Analysis

In order to effectively evaluate the performance of the solar cell surface defect detection algorithm proposed in this paper, the same solar cell surface defect PASCAL VOC2007 format dataset as in the literature [26] is used for the experiments. The dataset contains 1461 images with defects and 1 image without defects, and the image resolution is approximately 5232×2720. In this paper, the dataset is expanded to 2687 images with defects by slicing, of which 79 images contain only broken cell defects, 1760 images contain only crack defects, 379 images contain only unsoldered area defects, and 469 images contain mixed defects. The images in the dataset are randomly selected to form three sets, the training set consists of 2383 images, the validation set consists of 149 images, and the test set consists of 155 images. The specific composition of the datasets is shown in Table 1. The images in this dataset contain 215 broken cell defects, 4974 crack defects, and 961 unsoldered area defects, with a total of 6150 defects. The specific distribution of each type of defect in the three datasets is shown in Table 2.

In this experiment, Faster R-CNN is used as the baseline model, VGG16 and Resnet101 are used as the backbone network for training, and then the images in the test set are input to the model, and the categories and localization boxes of the defects in the test images are finally output by identifying and locating the three types of defects. The network parameters during training are set as follows: Batch_size is 6, Base lr is 0.001, weight decay and momentum factor are 0.0005 and 0.9, respectively, and the maximum number of iterations N is 100 [20]. The test environment is as follows: a 64-bit Linux system with a CPU of model i7-9700KF and a GPU of GTX2080 (11G video memory) for acceleration, and the algorithm is implemented using the pytorch deep learning platform. The comparison algorithms used in the experiments of this paper are: yolo v3 [14], yolo v5 [16], FPN [23], Cascade R-CNN [22], R-FCN [21], and Faster R-CNN (divided into VGG16 and Resnet101 as the backbone network) [20].

Table 1
The specific composition of the datasets

Datasets	Broken cell	Crack	Unsoldered area	Mixed	Total number
Training set	64	1571	319	429	2383
Validation set	5	94	30	20	149
Test set	10	95	30	20	155
Total number	79	1760	379	469	2687

Table 2
Distribution statistical table

Datasets	Broken cell	Crack	Unsoldered area	Total number
Training set	100	2312	403	2815
Validation set	15	261	62	338
Test set	100	2401	496	2997
Total number	215	4974	961	6150

4.1 Visual effect analysis

Figure 4 gives the enlarged images of the detection results of solar cell surface defects at the same location by Faster R-CNN and the proposed method, A and B are the enlarged images of the local detection results of Faster R-CNN with VGG16 as the backbone network and the proposed method, respectively, C and D

are the enlarged images of the local detection results of Faster R-CNN with Resnet101 as the backbone network and the proposed method, respectively. It can be seen that the detection result of Faster R-CNN with VGG16 as the backbone network in Fig. 4A(1) has the problem of misdetection, and the bounding box is larger compared with the defect size. The detection result of the proposed algorithm is shown in Fig. 4B(1), the position marked by the bounding box is more accurate under the condition of obtaining the same confidence. Comparing Fig. 4C(1) with Fig. 4. D(1), it can be found that the location marked by the bounding box of the detection result of the proposed algorithm is more accurate and the confidence is higher. Observing the second row of Fig. 4, it can be seen that the detection results of Faster R-CNN in Fig. 4A(2) and Fig. 4C(2) have the problems that the bounding box does not completely cover the defect location and the bounding box is large, respectively. Figure 4B(2) and Fig. 4D(2) show the detection results of the proposed algorithm, it can be seen that the location of the bounding box is more accurate. Comparing Fig. 4A(3) with Fig. 4B(3), it can be found that the detection results of the proposed algorithm have a slight advantage in predicting the location. In Fig. 4C(3), the detection result of Faster R-CNN with Resnet101 as the backbone network has the problem of missed detection, while the detection results of the proposed algorithm in Fig. 4D(3) accurately detect 3 crack defects, and the position marked by the bounding box is more accurate. On the whole, the proposed method has advantages in both the accuracy of the predicted location and the confidence of the detection results.

Figure 5 presents a comparative analysis of the bounding box regression accuracy for four types of defects, namely, broken cell defects, crack defects, unsoldered area defects and mixed defects. Comparing the bounding boxes for each type of defect obtained by yolo v3, yolo v5, FPN, Cascade R-CNN, R-FCN, Faster R-CNN and the proposed algorithm with the ground truth bounding boxes. As shown in Fig. 5, the real boxes are shown in the last column, which are indicated in blue. Figure 5a shows the local enlargement of the bounding box for the broken cells of different detection models. Comparing the real box, it can be concluded that yolo v3, FPN, Cascade R-CNN, and Faster R-CNN all detect broken cell defects, but the regression error is larger, and the IOU value between the bounding box of broken cell defect and the real box is lower, yolo v5 and R-FCN have higher detection accuracy, and the proposed algorithm has better performance, especially the proposed method with Resnet101 as the backbone network, which ensures higher regression accuracy. Figure 5b shows the local enlargement of the bounding box for cracks of different detection models, and it can be seen from the figure that Cascade R-CNN fails to detect the crack location accurately, while other methods detect the crack location, and the regression accuracy of R-FCN and the proposed algorithm is higher and the error is smaller. Figure 5c shows the partial enlargement of the bounding box for unsoldered areas of different detection models. It can be seen that the Cascade R-CNN bounding box is too small or fails to detect the unsoldered area location completely, and the FPN bounding box is too large, while the other methods detect the unsoldered area location more accurately, and the proposed algorithm has higher regression accuracy. Figure 5d shows the local enlargement of the bounding box for mixed defects of different detection models. Observing the figure, we can see that yolo v3, yolo v5 and Cascade R-CNN have high miss detection rate, R-FCN and the proposed algorithm have high regression accuracy, and the proposed algorithm has the best effect. In summary, the analysis shows that in the comparison of the regression accuracy of the four types of defective bounding boxes, the proposed algorithm can not only accurately detect the defects present, but also ensure a high regression accuracy of the bounding boxes, especially, the proposed algorithm with Resnet101 as the backbone network has the best effect.

4.2 Quantitative analysis

To further quantitatively compare the performance of the algorithms, this subsection uses yolo v3, yolo v5, Cascade R-CNN, R-FCN, FPN and Faster R-CNN as the compared algorithms for the experiments (the backbone networks used by each algorithm are shown in Table 3). Table 3 shows the defect detection results statistics of different algorithms on the same solar dataset. The detection accuracy of each algorithm on 677 graphs was randomly selected from the experimental results, and the algorithms were analyzed for the number of correct detections (Accurate), the number of false negative (FN), the number of false positive (FP), the accuracy of the three types of defects (Accuracy), false negative rate (FNR), and false positive rate (FPR). As can be seen from Table 3, in the detection results of broken cell defects, the FNR of the proposed algorithm based on both backbone networks are 0, and the Accuracy of the proposed algorithm based on Resnet101 network is the highest and the FPR is the lowest. In the detection results of crack defects, the Accuracy of the proposed algorithm based on both backbone networks are the highest and the FNR are the lowest, and the FPR of the proposed algorithm based on Resnet101 network is only 0.3% higher than that of Faster R-CNN with the best effect. In the detection results of unsoldered area defects, the proposed algorithms based on two backbone networks have the lowest FNR, and the proposed algorithm based on Resnet101 network has the highest Accuracy and the lowest FPR, which is the same as yolo v3, both of which are 0.97%.

In addition, this paper analyses the detection effects of different algorithms on the surface defects of the single solar cell image. yolo v3, yolo v5, Cascade R-CNN, R-FCN, FPN, Faster R-CNN and the proposed algorithm are used to detect the test set images, and the detection results are shown in Table 4. The mean average precision (mAP), which is an important index to evaluate the detection performance, is shown in the last column in Table 4. It can be seen that the proposed algorithm based on the Resnet101 backbone network has the highest mAP value, which is 6.69% higher than that of the yolo v3 algorithm with the lowest mAP value. Compared with Faster R-CNN, the mAP values of the proposed algorithm based on two backbone networks are increased by 2.89% and 0.83%, respectively. Overall, the proposed algorithm greatly improves the detection performance on the surface defectsof the solar cell.

Table 3
The defect detection statistical results of different defect models

	Methods	Yolo v3	Yolo v5	Cascade R-CNN	R-FCN	FPN	Faster R-CNN		Proposed Method	
	Backbone networks	DarkNet	Focus + CS	detnet59	Resnet101	Resnet101	Vgg16	Resnet101	Vgg16	Backbone networks
Broken cell	Accurate	69	79	67	76	78	75	85	Broken cell	Accurate
	FN	17	4	16	0	0	0	1	0	FN
	FP	4	7	7	14	12	15	4	10	FP
	Accuracy(%)	76.7	87.8	74.4	84.4	86.7	83.3	94.4	88.9	Accuracy(%)
	FNR(%)	18.9	4.4	17.8	0	0	0	1.1	0	FNR(%)
	FPR(%)	4.4	7.8	7.8	15.6	13.3	16.7	4.4	11.1	FPR(%)
Crack	Accurate	209	242	194	243	255	236	259	Crack	Accurate
	FN	62	22	75	14	18	12	16	6	FN
	FP	7	14	9	21	5	30	3	7	FP
	Accuracy(%)	75.2	87.0	69.8	87.4	91.7	94.9	93.2	95.3	Accuracy(%)
	FNR(%)	22.3	7.9	26.98	5.03	6.5	4.3	5.8	2.2	FNR(%)
	FPR(%)	2.5	5.03	3.24	7.55	1.8	10.8	1.1	2.5	FPR(%)
Unsoldered area	Accurate	252	286	266	271	283	252	292	Unsoldered area	Accurate
	FN	54	16	38	16	11	4	8	3	FN
	FP	3	7	5	22	15	53	9	23	FP
	Accuracy(%)	81.5	92.5	86.1	87.7	91.6	81.5	94.5	91.6	Accuracy(%)
	FNR(%)	17.5	5.2	12.3	5.2	3.6	1.3	2.6	0.97	FNR(%)
	FPR(%)	0.97	2.3	1.6	7.1	4.9	17.2	2.9	7.4	FPR(%)

Table 4
The mAP values of different defect model

Methods	Backbone networks	Brokencell	Crack	Unsoldered area	mAP(%)
Yolo v3	DarkNet	83.60	79.80	90.30	84.50
olo v5	Focus + CSP	88.00	81.60	98.50	89.40
Cascade R-CNN	detnet59	90.90	90.57	90.81	90.76
R-FCN	Resnet101	89.26	88.69	89.42	89.12
FPN	Resnet101	90.41	89.91	89.76	90.03
Faster R-CNN	Vgg16	86.34	84.21	89.17	86.57
	Resnet101	90.41	90.47	90.19	90.36
Ous Methods	Vgg16	90.06	88.74	89.59	89.46
	Resnet101	91.33	91.29	90.94	91.19

Table 5
Statistical table of test time of different modules

Methods	Total test time(s)	Average test time per image(s)
Faster R-CNN	520.113	0.768
Faster R-CNN + ACMFN	883.545	1.305
Faster R-CNN + TDC	550.403	0.813
Faster R-CNN + S-NMS	378.448	0.559
Ours Method	557.872	0.824

To evaluate the detection real-time performance of the proposed algorithm, the new network formed by Faster R-CNN combined with each module method in this paper and the overall network of the proposed algorithm are respectively tested on 677 test images. As shown in Table 5, the total test time and the average test time of each image are compared. It can be seen that, based on the Faster R-CNN method, the total test time increased by about 70% compared with the original method after the CCMFN is superimposed separately; the total test time is only about 6% longer than the original method after superimposing the TDC block alone; after superimposing the S-NMS alone, the total test time is only 73% of the original method due to the reduction of a large number of redundant prediction boxes. However, after all the above modules are superimposed in proposed method, the average test time per image is only increased by 0.056s compared with the original method. It can be seen that the proposed algorithm meets the real-time detection requirements.

4.3. Ablation experiment

To verify the feasibility and effectiveness of the various strategies adopted in the proposed algorithm, ablation experiments are carried out. Figure 6 verify the effectiveness of the S-NMS strategy from visual perspectives. Figure 6 compares the visual effects of the partially enlarged images of the bounding box for many defects before and after using the S-NMS strategy in the Faster R-CNN network. By comparing the two rows of enlarged images, it can be clearly seen that the redundant boxes and false detection boxes are both reduced after adding the S-NMS strategy, and the prediction confidence is further improved while retaining the more accurate detection boxes. It can be concluded that the S-NMS strategy used in the proposed algorithm effectively improves the detection accuracy of solar cell surface defects.

Table 6
The mAP values statistics of different network module

Backbone networks	Vgg16								Resnet101							
ACMFN			√	√			√	√			√	√			√	√
TDC					√	√	√	√					√	√	√	√
S-NMS		√		√		√		√		√		√		√		√
mAP(%)	86.19	86.57	88.43	88.60	88.37	88.59	89.29	89.46	90.36	90.40	90.68	90.71	90.59	90.63	91.17	91.19

In order to further verify the effectiveness of the three modules of CCMFN, TDC and S-NMS, the different modules are added to the Faster R-CNN model with VGG16 and Resnet101 as the backbone networks, respectively, and their mAP values are calculated in this paper, and the experimental results are shown in Table 6. It can be seen from the mAP value statistics table that the mAP values of the models based on the two backbone networks have been improved to varying degrees when the above three modules are added individually or jointly. When the modules are added individually, the largest mAP value improvement was obtained for the CCMFN module and the smallest improvement was obtained for the S-NMS module. when the three modules were used simultaneously, the largest mAP value improvement was obtained with an improvement of 3.27% and 0.83%, respectively. The above analysis shows that each module plays an important role in the improvement of network performance, and the three modules used together have the best effect, which verifies the feasibility and effectiveness of the three strategies adopted in the proposed algorithm.

5 Conclusion

Aiming at the difficulty of detecting solar cell surface defects due to the small defects, various forms, and large scale changes, this paper extracts multi-scale solar defect target features from different feature layers based on the Faster R-CNN. The convolutional layers with different receptive fields are further used for feature extraction and fusion. The NMS evaluation criteria is redesigned, and the cosine similarity is introduced into the candidate box aspect-ratio similarity, which greatly improves the detection performance of the network. The experimental results show that, compared with the latest deep learning target detection models, the proposed model not only has higher detection accuracy, but also has lower probability of false detection and missed detection in various types of defects detection.

Declarations

Ethics approval Not applicable

Competing interests The authors declare no conflicts of interests.

Authors' contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Yanling Wang, Ting Hou, Xiong Zhang, Hong Shangguan, Pengcheng Zhang, Jie Li, and Baoyue Wei. The first draft of the manuscript was written by Yanling Wang and Ting Hou. All authors reviewed and approved the manuscript.

Availability of data and materials

All authors are sure that all data and materials as well as software application support published claims and comply with field standards.

References

1. Li, Q., Wang, W., Ma, C., Zhu, Z.: Detection of physical defects in solar cells by hyper spectral imaging technology. Opt. Laser Technol. 42, 1010–1013 (2010)
2. Drabczyk K., Kulesza-Matlak G., Drygała A., et al.: Electroluminescence imaging for determining the influence of metallization parameters for solar cell metal contacts. Solar Energy 126, 14–21 (2016)
3. Sawyer, David, E., Kessler, Herbert, K.: Laser scanning of solar cells for the display of cell operating characteristics and detection of cell defects. IEEE Transactions on Electron Devices 27(4), 864–872 (1980)
4. Hilmersson, C., Hess, D.P., Dallas, W., Ostapenko, S.: Crack detection in single-crystalline silicon wafers using impact testing. Applied Acoustics 69(8), 755–760 (2008)
5. Jen, C.Y., Richter, C.: Doping profile recognition applied to silicon photovoltaic cells using terahertz time-domain spectroscopy. IEEE Transactions on Terahertz Science and Technology 4(5), 560–567 (2014)
6. Tsai, D.M., Li, G.N., Li, W.C., et al.: Defect detection in multi-crystal solar cells using clustering with uniformity measures, Adv. Eng. Inf. 29 (3), 419–430 (2015)
7. Qian, X., Zhang, H., Yang, C., et al.: Micro-cracks detection of multicrystalline solar cell surface based on self-learning features and low-rank matrix recovery, SensorRev. 38(3), 360–368 (2018)
8. Li, W.C., Tsai, D.M.: Wavelet-based defect detection in solar wafer images with inhomogeneous texture, Pattern Recogn. 45(2), 742–756 (2012)
9. LeCun, Y,, Bengio, Y, Hinton G.: Deep learning. Nature 52l (7553), 436–444 (2015)
10. Liu, L., Ouyang, W., Wang, X., et al.: Deep learning for generic object detection: A survey. International journal of computer vision 128(2), 261–318 (2020)
11. Girshick, R., Donahue, J., Darrell, T., et al.: Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition 580–587 (2014)
12. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition 779–788 (2016)
13. Redmon, J., Farhadi, A.: Yolo9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition 7263–7271 (2017)
14. Redmon, J., Farhadi, A.: Yolov3: An incremental improvement. arXiv preprint arXiv: 1804. 02767 (2018)
15. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv: 2004. 10934 (2020)
16. Zhu, X., Lyu, S., Wang, X., et al.: TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2778–2788 (2021)
17. Liu, W., Anguelov, D., Erhan, D., et al.: Ssd: Single shot multibox detector. European conference on computer vision. Springer, Cham 21–37 (2016)
18. He, K., Zhang, X., Ren, S., et al.: Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE transactions on pattern analysis and machine intelligence 37(9), 1904–1916 (2015)
19. Girshick, R., Iandola, F., Darrell, T., et al.: Deformable part models are convolutional neural networks. Proceedings of the IEEE conference on Computer Vision and Pattern Recognition 437–446 (2015)
20. Ren, S., He, K., Girshick, R., et al.: Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems 91–99 (2015)

21. Dai, J., Li, Y., He, K., et al.: R-fcn: Object detection via region-based fully convolutional networks. Advances in neural information processing systems 29 (2016)

22. Cai, Z., Vasconcelos, N.: Cascade R-CNN: Delving into high quality object detection. Proceedings of the IEEE conference on computer vision and pattern recognition 6154–6162 (2018)

23. Lin, T.Y., Dollár, Piotr, Girshick R., et al.: Feature pyramid networks for object detection. Proceedings of the IEEE conference on computer vision and pattern recognition 2117–2125 (2017)

24. Bartler, A., Mauch, L., Yang B., et al.: Automated detection of solar cell defects with deep learning. 2018 26th European Signal Processing Conference (EUSIPCO), IEEE 2035–2039 (2018)

25. Chen, H., Pang, Y., Hu, Q., et al.: Solar cell surface defect inspection based on multispectral convolutional neural network. Journal of Intelligent Manufacturing 31(2), 453–468 (2020)

26. Zhang, X., Hao, Y., Shangguan, H., et al.: Detection of surface defects on solar cells by fusing Multi-channel convolution neural networks. Infrared Physics & Technology 108, 103334 (2020)

27. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition 770–778 (2016)

28. Li, Y., Chen, Y., Wang, N., et al.: Scale-aware trident networks for object detection. Proceedings of the IEEE/CVF International Conference on Computer Vision 6054–6063 (2019)

29. He, Y., Zhang, X., Savvides, M., et al.: Softer-nms: Rethinking bounding box regression for accurate object detection. arXiv preprint arXiv:1809.08545 2(3), 69–80 (2018)
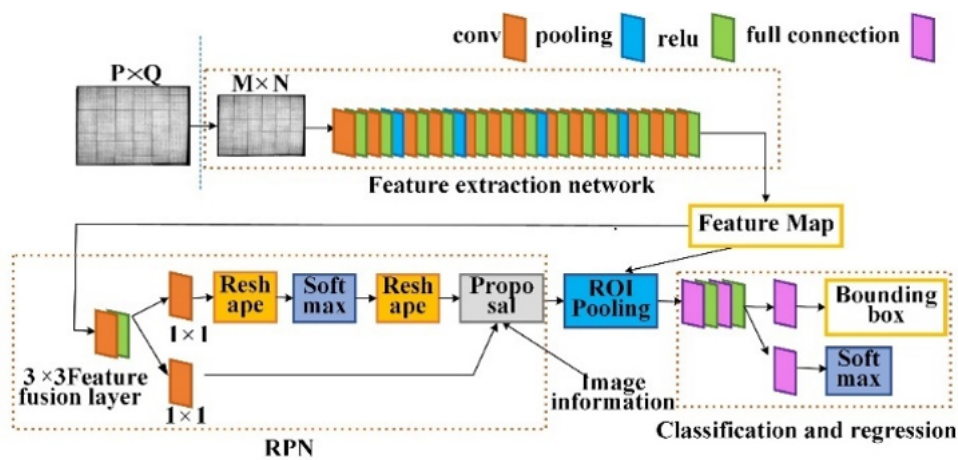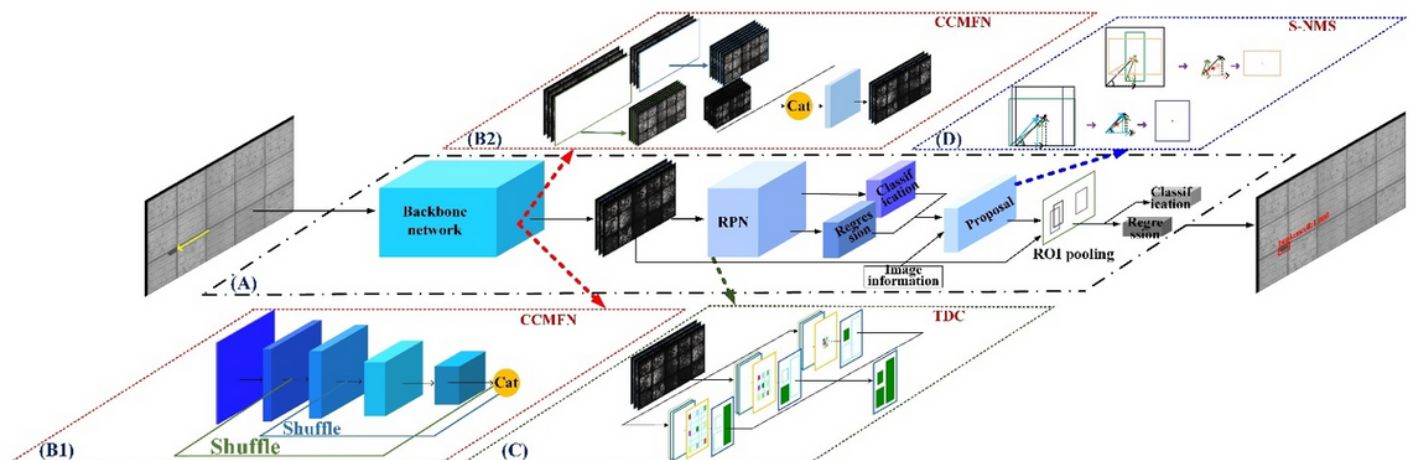
Figures

Figure 1

Faster R-CNN structure
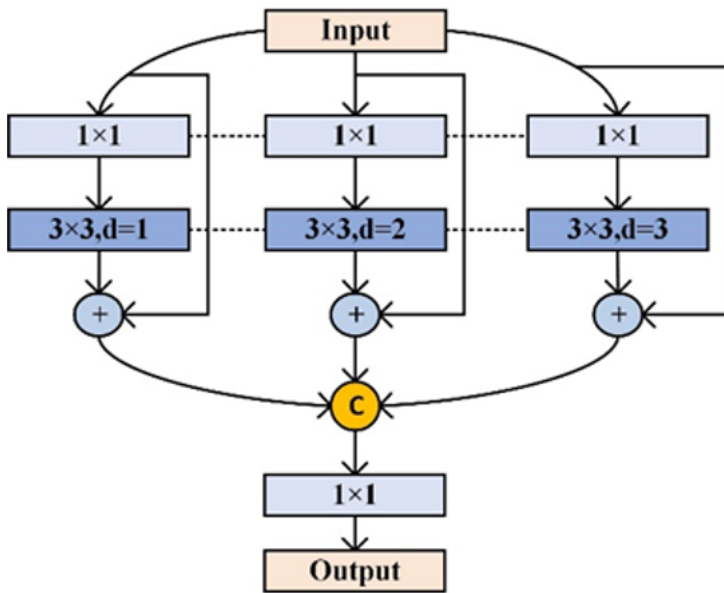
Figure 2

Overall network structure

Figure 3
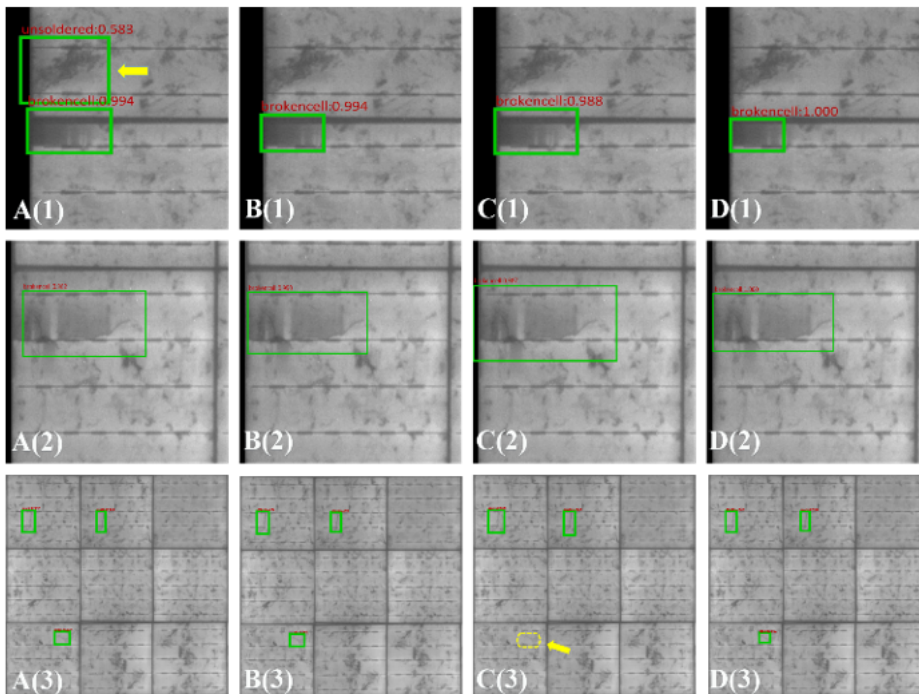
Three-branch Dilated Convolution (TDC)

Figure 4

Partial enlarged images of the defect detection results obtained using Faster R-CNN and our method
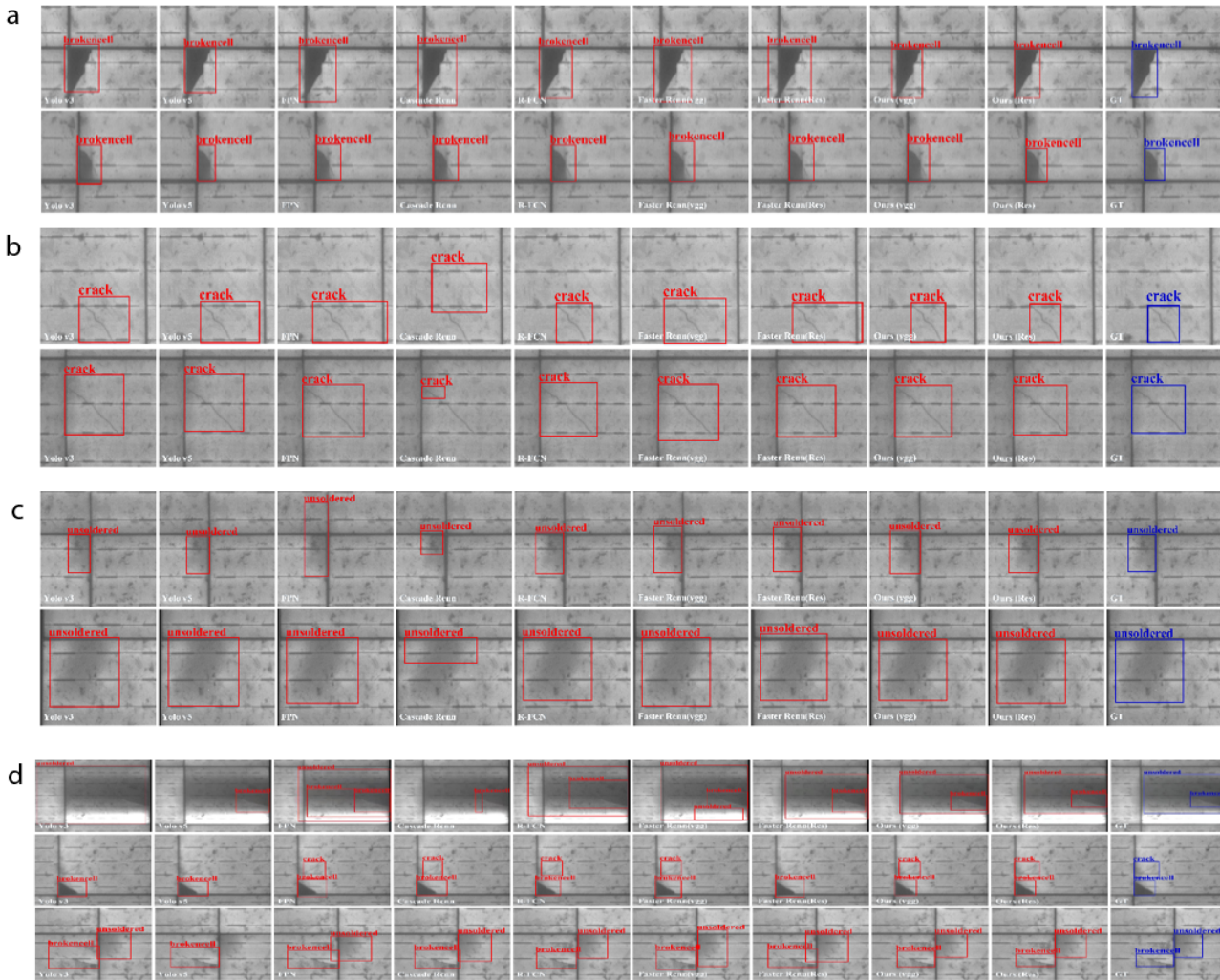
Figure 5

a Partially enlarged images of the bounding box for brokencells of different defect models

b Partially enlarged images of the bounding box for cracks of different defect models

c Partially enlarged images of the bounding box for uns-oldereds of different defect models

d Partially enlarged images of the bounding box for mixed defects of different defect models
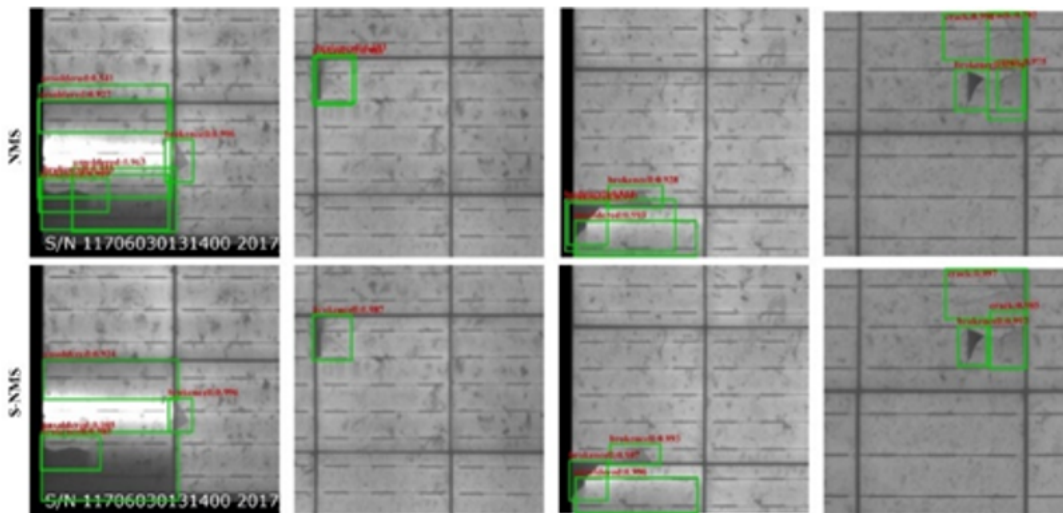

Figure 6

Partially enlarged images of the bounding box for many defects of NMS and S-NMS