IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 24 June 2023, accepted 8 July 2023, date of publication 11 July 2023, date of current version 18 July 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3294344

RESEARCH ARTICLE

Solar Cell Surface Defect Detection Based on Optimized YOLOv5

SHA LU, KAIXING WU, AND JINXIN CHEN
School of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, China
Corresponding author: Sha Lu (lusha123123@163.com)

This work was supported in part by the Science and Technology (S&T) Program of Hebei Province under Grant 20475702D.

ABSTRACT Traditional vision methods for solar cell defect detection have problems such as low accuracy and few types of detection, so this paper proposes an optimized YOLOv5 model for more accurate and comprehensive identification of defects in solar cells. The model firstly integrates five data enhancement methods, namely Mosaic, Mixup, hsv transform, scale transform and flip, to expand the existing data set to improve the feature training accuracy and enhance the robustness of the model; secondly, CA attention mechanism is introduced to improve the feature extraction ability of the model; to address the problems of different target defect classification and localization concerns, the detection head in the original model is replaced with a decoupling head, which significantly improve the detection accuracy of the model without affecting the convergence speed of the model. The results show that the optimized model achieves an mAP of 96.1% on the publicly available dichotomous ELPV dataset, and can identify and locate a variety of common defects in the PVEL-AD dataset, while the mAP can reach 87.4%, an improvement of 10.38% compared with the original YOLOv5 model, which enables the model to perform more accurately while ensuring the real-time requirement of solar cell surface defects detection task.

INDEX TERMS YOLOv5, solar cell, defect detection, coordinate attention, decoupled detection head.

I. INTRODUCTION

With the proposal of "carbon peaking" and "carbon neutrality", the use of clean energy has attracted more and more attention. Among them, solar energy stands out from many clean energy sources with its advantages of safety, reliability, low cost, and wide application range. At present, the use of solar energy is mainly realized by converting solar energy into electrical energy through silicon cells. However, due to the sensitivity of its raw materials, the product is prone to cracks, short circuits and black cores during production process. Therefore, identifying the defects in the solar cell production process through efficient detection means can greatly improve the yield rate of solar cells, which is particularly important for improving the photoelectric conversion efficiency.

Traditionally, the detection of surface defects of solar cells mainly adopts manual visual inspection and machine vision detection based on industrial cameras [1], however, these methods not only have a large workload and low efficiency, but also are easily affected by the subjective factors of the inspectors, resulting in missed and wrong detections. Since 1950, with the introduction of the concept of deep learning, deep learning models represented by convolutional neural networks have been widely used in the fields of image recognition [2], [3] and natural language processing [4], [5], but they cannot be directly applied to the detection of solar cell defects. Mainly because (1) it is difficult to collect solar cell defect images, and the amount of available data is small; (2) there are many types of defects in solar cells, and the shapes are different; (3) solar cell defect detection is susceptible to background interference; (4)with the continuous training and downsampling of the network, many small defect features gradually disappear. These issues are all challenges to be faced when detecting solar cell defects.

Recently, the networks used for target detection include one-stage network SSD [6], YOLO and two-stage network R-CNN, Faster R-CNN [7], etc. Among them, R-CNN and

The associate editor coordinating the review of this manuscript and approving it for publication was Zhongyi Guo.

Faster R-CNN have high detection accuracy, but they are slow, take up a lot of memory, and require a lot of computational resources. SSD for small target detection still needs to be improved and can not be detected in real time like YOLO. YOLO series models have greater advantages over several other mainstream methods in terms of detection accuracy and detection speed. The YOLOv5 is not only the relatively small model and superior performance among YOLO series, but also achieves the best balance between accuracy and speed of detection in many application scenarios. Especially in recent years, different researchers have improved and optimized YOLOv5 according to their needs, and the improved YOLOv5 has performed very well in many fields.

After the above analysis and demonstration, the one-stage network YOLOv5 plays an important role in target detection has a powerful detection real-time processing capability and low hardware requirements for real-time detection. Based on this, this paper proposes an optimized YOLOv5 model for the complexity and specificity of solar cell defects, which can identify and detect various defects of solar cells more accurately. The main contributions of this paper can be summarized as follows:

1.Using the idea of combining Mosaic, Mixup, hsv transformation, scale transformation, and flipping five kinds of data enhancement, the data set is expanded without losing the original feature information, and the robustness of the model is enhanced;

2.Adding CA attention mechanism between the neck and head of the model to enhance the model's ability to select important channel information and make the model's localization and target recognition ability more accurate.

3.The decoupling head is used to replace the original YOLOv5 detection head to solve the problem of conflict between classification and positioning due to different concerns, thereby improving the accuracy of detection.

The remainder of this paper is organized as follows, Part II introduces some related work on detecting solar cell defects; Part III mainly introduces the methods used in this paper; Part IV analyzes the model method of this paper through comparative experiments and ablation experiments; Part V elaborates in conclusion.

II. RELATED WORK

In recent years, detection methods based on machine vision and computer vision have been continuously applied to the detection of surface defects of solar cells [8]. In 2019, the literature [9] designed a solar panel crack detection device based on deep learning algorithms in Halcon image processing software for the most common defects in the solar panel production process, which can effectively detect cracks in solar panels, reduce the rate of late defective products, improve the production quality of solar cells, and reduce energy waste and labor costs. However, the detection accuracy and detection type of the method proposed in this literature need to be improved. In the same year, the literature [10], edge detection and Hough transform based image processing techniques were adapted for efficient identification of faults. The processed image is subjected to feature extraction and passed through a classification algorithm for localization and identification of the type of fault. Although the detection accuracy has been improved, the computation is large and the detection speed does not meet the requirement of real-time. In 2020, a novel structure-aware-based crack defect detection scheme (SACDDS) is proposed. Experimental results show that the proposed method can completely extract crack defect in the inhomogeneously textured background, which is well effective and outperforms the previous methods [11]. However, the method only investigates the detection of one defect type of cracks. It cannot meet the actual production requirements of solar cell defect detection. In the literature [12], an improved convolutional neural network is proposed for the detection of defects in inspection panels, and the model can achieve the recognition and detection of multiple defects such as broken grids, open welds, and hidden cracks at the same time, but its detection accuracy has more room for improvement. In 2021, Zubair et al. [13] of the University of New South Wales in Australia used a target classification neural network to detect defects in solar cell PL images, but the network can only classify solar cell defects, and cannot accurately locate solar cell defects. Wang et al. [14] of East China Normal University proposed an automatic detection and classification method for solar cell defects based on EL imaging. The team developed an unsupervised algorithm based on a recurrent neural network, which can automatically detect defects based on EL images. Defects and Classification. This method is the first attempt to combine automatic defect detection with defect texture classification. According to the experimental results of various types of solar cells, the average uncertainty of this method can reach as low as 5.15%, and the optimization rate can reach as high as 98.9%. Although this method can save a lot of time and cost waste caused by sample labeling, it is difficult to meet the classification, recognition and detection requirements for complex problems due to the randomness of unsupervised learning. In 2022, Alaa et al. [15] proposed an improved anomaly detection method for EL imaging of PV cell surface defects based on Faster R-CNN, which integrates lightweight channels and spatial convolutional attention modules. It can analyze crack defects in complex scenes more effectively. In the same year, a series of improvements were made to the YOLOv5 model in the literature, and the accuracy of the improved model for detecting solar cell defects was substantially improved [16]. However, both methods are only for three easy to detect defect types Finger, Black core, Crack, and for other common types of detection still need to be studied. In 2023, the literature [17] proposed a photovoltaic panel defect detection method based on YOLOv5's tiny target prediction head (GBH-YOLOv5). Lightweight and detection types also have limitations. An efficient Real-Time Multi Variant Deep learning Model (RMVDM) is presented in this

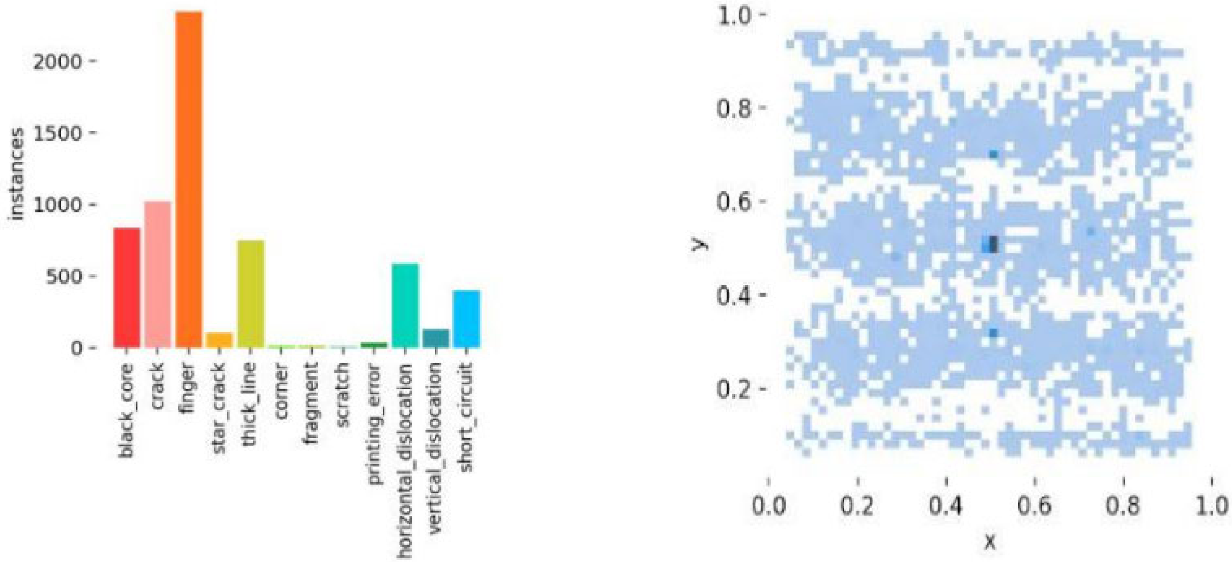


FIGURE 1. The number and location distribution of various defects.

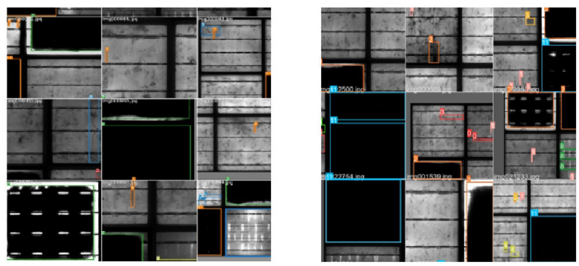
article to handle this issue. The proposed model produces efficient results with around 97% in defect detection and localization with higher accuracy and less time complexity [18]. However, this method uses a small data set and can be fitted or overfitted. In the same year, literature [19] was optimized on the basis of the Faster R-CNN model, which combines lightweight channels and spatial convolution attention modules, which can effectively identify and analyze crack defects in solar cell complex data sets. At the same time, the added clustering and loss functions are used to improve the model's ability to detect small target defects, but the model only detects three common defects: cracks, black corners, and broken grids.

According to the above research, it is found that there are still many deficiencies in the detection of solar cell defects.For example, 1) Since it is difficult to extract solar cell defect images, the small amount of data is easy to cause insufficient training and low accuracy; 2) There are many types of solar cell defects, and the same type of defect has different shapes, which is easy to misdetect and miss; 3) The existing defect detection model of solar cells still has problems such as poor recognition ability of target defects and complex feature defects. These issues will affect the reliability of industrial production, therefore, further research is needed.

III. METHOD

A. YOLOv5 OVERVIEW

There are four versions of YOLOv5, namely YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x [20]. The four models have different depth and width parameters. The smaller the network model, the lower the hardware requirements and the easier it is to deploy. Therefore, this paper uses YOLOv5s



(a) Mosaic data enhancement (b) Hybrid data enhancement

FIGURE 2. Training effect before and after data enhancement.

with the smallest network depth and width as the basic model, and deepens and expands on this basis. YOLOv5 can be roughly divided into three parts: Backbone, Neck, and Head. Among them, Backbone adopts CSPDarknet53 architecture; Neck is composed of Feature Pyramid Network (FPN) [21] and Path Aggregation Network (Path Aggregation Network, PAN) [22]; Head has a total of three detection heads.

B. HYBRID DATA AUGMENTATION

Figure 1 shows the number and location distribution of various types of defects in the data set PVEL-AD used in this paper. It can be seen from the information in the figure that the number of some defects is very sufficient, reaching more than 2000, but the number of some defects is very small, even less than 10. In order to improve the overall detection accuracy, it is ensured that each defect has enough sample size, improve the ability of the neural network to identify each defect category, this paper uses Mosaic data enhancement, Mixup data enhancement [23], hsv transformation, scale transformation, and flipping method to process according to the defect type

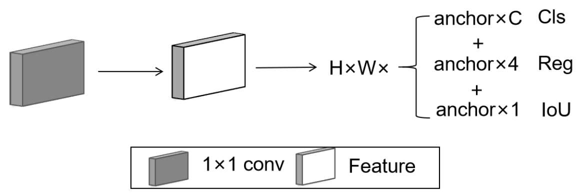


FIGURE 3. **YOLOv5 coupled detection head.**

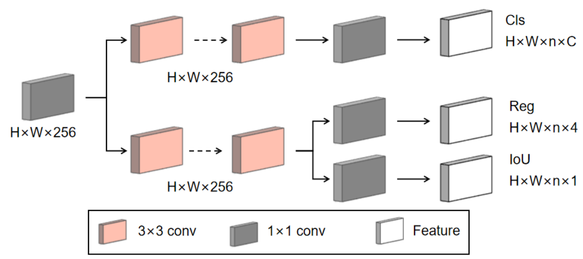


FIGURE 4. **Schematic diagram of decoupled detection head.**

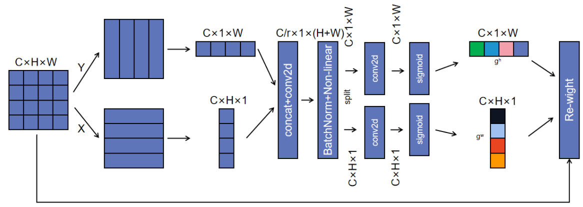


FIGURE 5. **Structural diagram of CA attention mechanism.**

and actual demand of solar cells The data set increases the number of samples of various defects to improve the training accuracy of the model and enhance the robustness of the model.

Mosaic data enhancement is to randomly select four images from the original data set for splicing, and then crop them into an image of the same size as the original image to complete the data enhancement.

As the network depth continues to increase, the sensitivity of the network model to adversarial examples will decrease. Therefore, this paper also uses the Mixup data enhancement method to fuse two different images to complete data enhancement to alleviate this problem. At the same time, this method can also reduce the memory of wrong labels, thereby enhancing the robustness of the model.

The hsv transformation can be enhanced from three angles of hue, saturation, and brightness. It improves the richness of the data set without destroying the key information in the image, so that the YOLOv5 model can see more data during the training process, which greatly improves the It reduces the computational cost and is a very practical data enhancement method.

In this paper, we adopt the idea of combining five kinds of data enhancement, and it can be seen from Figure 2 that the optimized YOLOv5 can increase the number of training samples by mixing data enhancement during the training process, which solves the problem of easy fitting and overfitting due to the small number of training samples for defect detection to a certain extent and enhances the robustness of the model.

C. DECOUPLED DETECTION HEAD

In object detection, the conflict between classification tasks and regression tasks is a well-known problem [24]. The reason for the conflict is that classification and localization focus on different points, among which classification focuses on the texture content of the target, while localization focuses on the edge information of the target. Therefore, if the same feature map is used for classification and positioning, the detection effect will be greatly reduced. However, in this paper, YOLOv5 is used to detect solar cell defects, not only for the defect localization task, but also for the defect classification task. As shown in Figure 3, the detection head of YOLOv5 is implemented directly through a 1×1 convolutional layer, which is a coupled detection head with no separation of the tasks of classification and regression, So the YOLOv5 model also has the problem of classification and localization conflicts. Wu et al. analyzed the two subtasks of classification and positioning in the detection task for this problem [25], and found that fc-head is more suitable for classification tasks, and conv-head is more suitable for positioning tasks. With the help of this idea, this paper improves YOLOv5 by replacing a new decoupling head and using different branches to perform calculations, which can not only speed up the convergence speed, but also improve the detection accuracy. The specific structure is shown in Figure 4.

This decoupled detection head first passes through a convolution module with a convolution kernel size of 1×1, a layout of 1, padding of 0, and a convolution kernel number of 256. This convolution module contains convolution Multilayer, BN and SiLu activation functions, through this convolution module, the number of channels can be unified to 256. Then two branches are parallelized. Both branches use a convolution module with a convolution kernel size of 3×3, a layout of 1, a padding of 1, and a number of convolution kernels of 256. Then the first branch is connected with a 1×1 convolutional layer to obtain a branch (cls branch) predicted by this paper for the target category information, and the other branch parallels two 1×1 convolutional layers to obtain a predicted target A regression parameter branch(Reg branch), a prediction branch (obj branch). The two branches complete the defect localization and classification tasks separately, solving the problem of conflicting localization and classification tasks, thus improving the detection performance of the model.

D. ATTENTION MECHANISM

In solar cell defect images, in addition to a large amount of defect information, there is also a large amount of complex background information. These complex background information will undergo multiple iterations during the

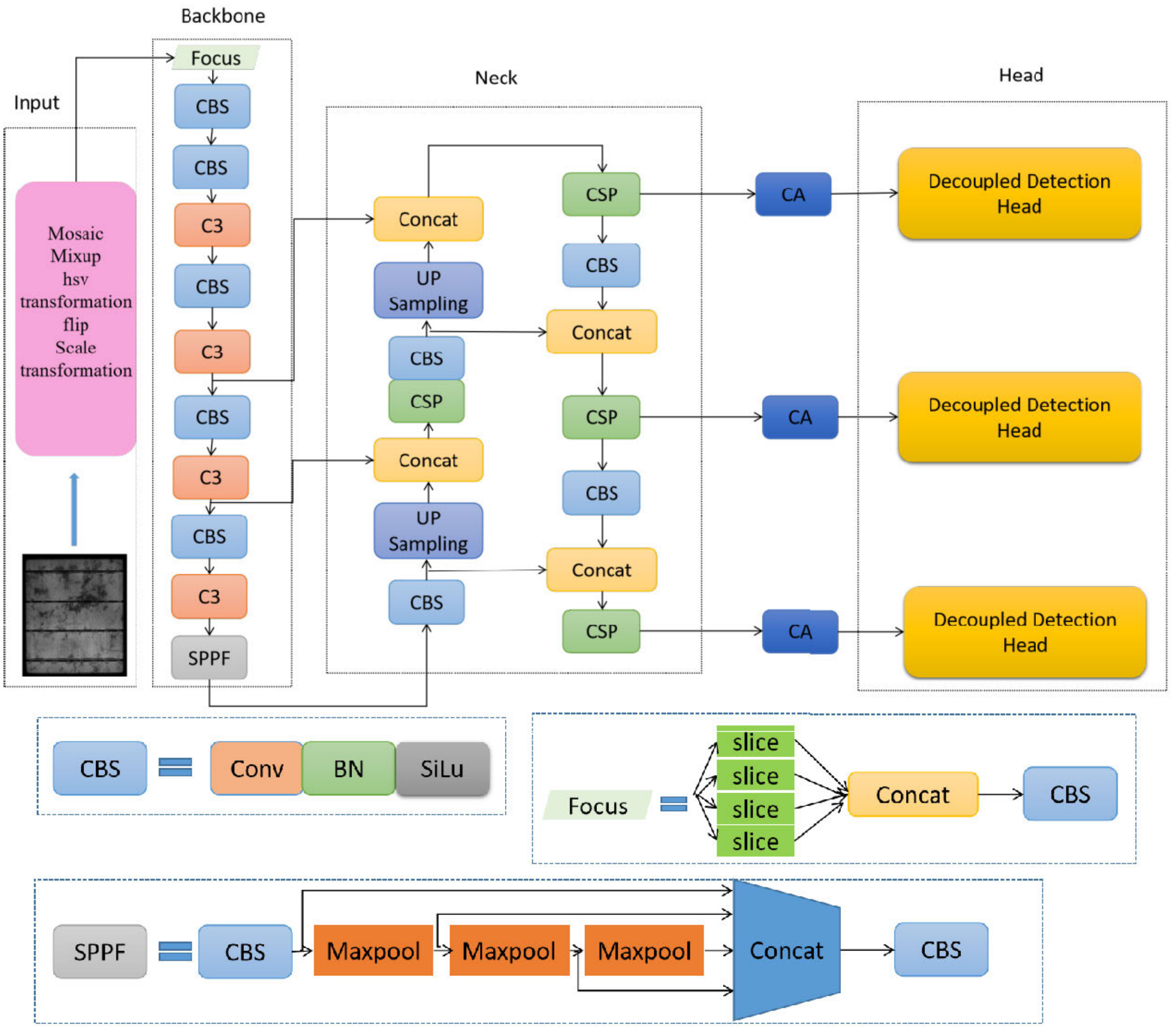


FIGURE 6. Optimized YOLOv5s structure schematic.

convolution operation, resulting in a lot of interference information, which affects the accuracy of detection. To solve this problem, it is a good choice to introduce attention mechanism. Currently, most of the attention mechanisms of lightweight networks use SE [26] modules, but SE only considers inter-channel information and ignores location information. Although the later BAM [27] and CBAM [28] try to extract location attention information by convolution after reducing the number of channels, convolution can only extract local relations and lacks the ability of long-distance relation extraction. But in 2021, an efficient attention mechanism CA (coordinate attention) was proposed in the literature [29], which can encode both horizontal and vertical location information into channel attention, enabling mobile networks to focus on a large range of location information without excessive computational effort. At the same time, the CA attention mechanism not only acquires inter-channel information, but also considers direction-related location information, which helps the model to better locate and identify targets, and is flexible and lightweight enough to be simply inserted into the core structure of the mobile network. Therefore, this paper introduces the coordinate attention mechanism CA in the YOLOv5 model and adds it between the neck and head of YOLOv5 so that the network model can focus on a large range of location information without too much computation. Thus, the model in this paper has better detection performance.

The CA module uses two pooling kernels with sizes (H×1) and (1×W) respectively to encode one-dimensional features in the height and width of the input feature map, thus obtaining two Feature map output in direction:

$$Z_c^h(h) = \frac{1}{w}\sum_{i=0}^{w} x_c(h, i) \tag{1}$$

$$Z_c^w(w) = \frac{1}{H}\sum_{j=0}^{H} x_c(j, w) \tag{2}$$

Among them, C represents the number of channels of the input image, and H and W represent the height and width of the input image.

Then send the output of the two feature maps to the convolution function with a weight of 1×1. When the dimension of the feature map drops to C/r, batch normalization can be performed. Finally, after the Sigmoid activation function, you will get The feature map f has a size of 1×(W+H)×C/r, where r is used to control the ratio of channel downsampling in convolution.

$$f = \sigma\left(F_1\left(Z_c^h, Z_c^w\right)\right) \tag{3}$$

The feature map f will be divided into two separate feature vectors on the channel, and can be transformed by two 1×1 convolutions and channel conversions respectively. The converted size and input have the same number of channels C×H×1 and C×1×W. After calculating activation function Sigmoid, two attention weight graphs are obtained

$$g^h = \sigma\left(F_h\left(f^h\right)\right) \tag{4}$$

$$g^w = \sigma\left(F_w\left(f^w\right)\right) \tag{5}$$

Finally, after multiplying the original input feature map with the two weight maps, the final output map is obtained:

$$y_c(i,j) = x_c(i,j) \times g^h(i) \times g^w(j) \tag{6}$$

IV. EXPERIMENT

A. DATA SET

This paper employed two currently publicly available solar cell defect datasets. The first dataset is ELPV [30]created and made public by Buerhop et al, This dataset contains 2624 samples of 300 × 300 pixel 8-bit grayscale images extracted from 44 solar modules with different degrees of functional degradation and defects, but this dataset does not provide images with defect markers, only the data are classified according to the degree of defects as 0% defect-free proportion, 33% probable defect proportion, 66% probable defect proportion, and defect proportion 100% four types are shown in Figure 7.

The second dataset is jointly released by Hebei University of Technology and Beijing University of Aeronautics and Astronautics the PVEL-AD dataset [31]. The dataset contains 1 class of non-anomalous images and images of abnormal defects with 12 different classes, such as cracks (lines and stars), broken grids, black cores, misalignment, thick lines, scratches, chips, broken Corners and material defects. The types of defects are shown in Figure 8.

B. EXPERIMENTAL CONDITIONS

The experimental environment of this research is Windows 10 operating system. The neural network framework uses Pytorch1.11.0 to build the experimental platform, using GPU RTX3090, CPUi5 128g, cuda version 11.3, and python language environment 3.7. Batch size is set to 16, and Epoch is set to 300 times.

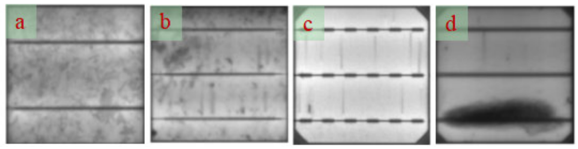


FIGURE 7. Samples of EL images of solar cells from ELPV dataset: a) defect certainty = 0, b) defect certainty = 0.33, c) defect certainty = 0.66, d) defect certainty = 1.

C. EVALUATION INDICATORS

The performance of the model is evaluated by Precision, Recall and mean average precision (mAP), and the formula for each metric is shown below:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{7}$$

$$\text{Re call} = \frac{TP}{TP + FN} \tag{8}$$

$$\text{AP} = \int_0^1 \text{PdR}. \tag{9}$$

$$\text{mAP} = \frac{\sum_{i=1}^{N} \text{AP}_i}{N} \tag{10}$$

In the formula: Precision is the proportion of correct prediction results; Recall is the proportion of all targets that are correctly predicted; TP (true positive) indicates the number of defects detected in the defective image, TN (true negative) indicates the number of defects detected in the defect-free image, FN (false negative) indicates the number of defects detected in the defect-free image, and FP (false positive) indicates the the number of defects detected in the defect-free image. The AP value is the area of the P-R curve, and mAP is the average value of all the corresponding solar cell surface defect categories.

D. EXPERIMENTAL RESULTS AND ANALYSIS

1) ABLATION EXPERIMENT

In order to better and more comprehensively analyze the performance impact of each module in the solar cell surface defect detection model built in this paper and to verify the effectiveness of each module structure, ablation experiments were designed and trained in this paper, and the test results are shown in Table 1: Each module and each optimization step introduced in this paper effectively improves the accuracy of the model, where the mAP is improved by 10.38% by incorporating hybrid data enhancement, CA attention mechanism and decoupling head. The results show that this model has been accumulated by various improvement methods, the model accuracy has been gradually improved, and the model has achieved a better detection effect.

2) ANALYSIS OF DETECTION RESULTS

After the model training is completed, the improved model is evaluated by plotting the corresponding curves with appropriate smoothing through the metrics within the training log. As shown in Figure 9, the comparison of the bounding box

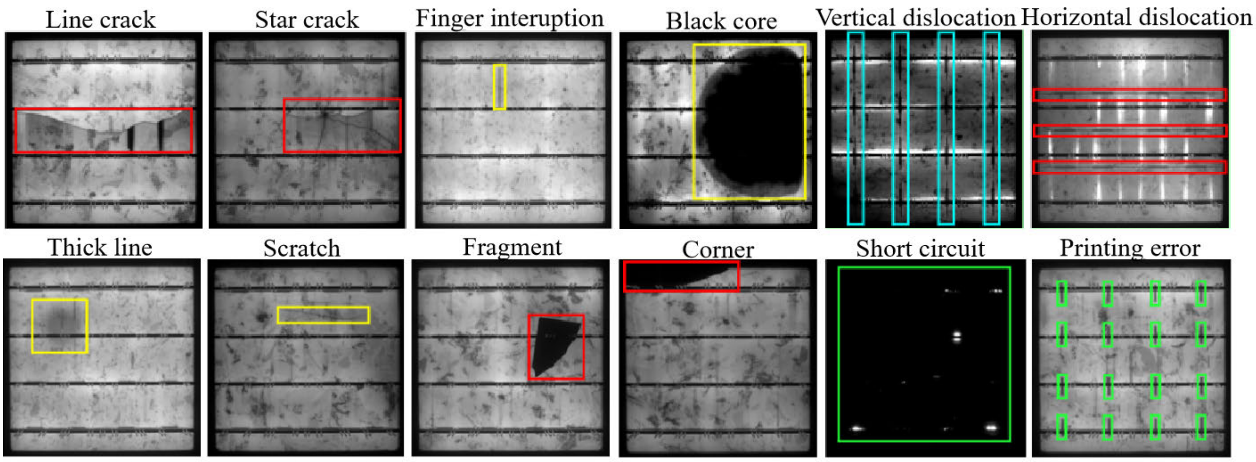


FIGURE 8. The morphology of 12 types of abnormal defects in the PVEL-AD dataset.

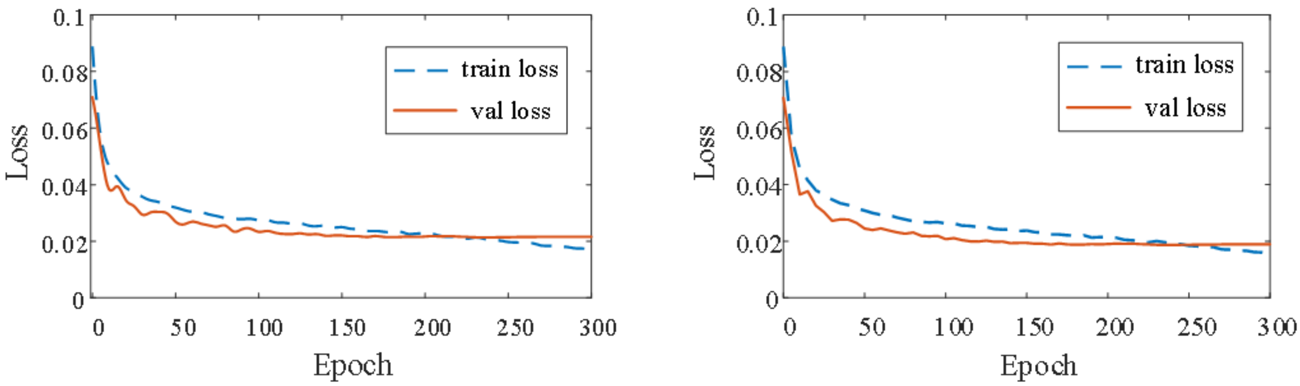


FIGURE 9. Loss plots of two YOLOv5s detection models.

TABLE 1. Statistical results of ablation experiments.

Original	Hybrid Data Enhancement	CA	Decoupled Detection Head	mAP @0.5(%)	mAP Increase
✓				77.02	
✓	✓			85.40	8.38
✓	✓	✓		87.00	1.60
✓	✓	✓	✓	87.40	0.40

TABLE 2. Performance comparison results of different models.

Models	P%	R%	mAP@0.5(%)
YOLOv5s	90.60	78.13	77.02
SSD	79.80	89.42	75.36
Faster R-CNN	87.26	80.98	78.60
BAF-Detector	\	\	80.77
ours	91.34	91.32	87.40

loss values of the optimized network and the original network is shown.

As can be seen from Figure 9, the loss values of the optimized YOLOv5 model for both the training and validation sets drop faster at the beginning of training. As the number of iterations increases, it gradually smoothes out and the inflection point appears earlier than the original YOLOv5 model. Finally, the loss values of the optimized YOLOv5 model for the training and validation sets are stable at around 0.0186 and 0.0197, respectively, which are lower than those of the original YOLOv5 at around 0.0189 and 0.0202 for the training and validation sets. The results show that the optimized YOLOv5 model has faster convergence speed and higher regression accuracy.

Figure 10 shows the mAP@0.5 curve of the original YOLOv5 and optimized model training. As can be seen from the figure, the original YOLOv5 has large ups and downs on the first 100 Epoch iterations, but the optimized YOLOv5 becomes smoother at around 60 Epochs. The overall optimized YOLOv5 curve is always higher than the original YOLOv5 curve, which means that the detection accuracy of the optimized YOLOv5 model is generally higher than that of the original YOLOv5 model, and the training curve of the optimized YOLOv5 model is smoother, indicating that the

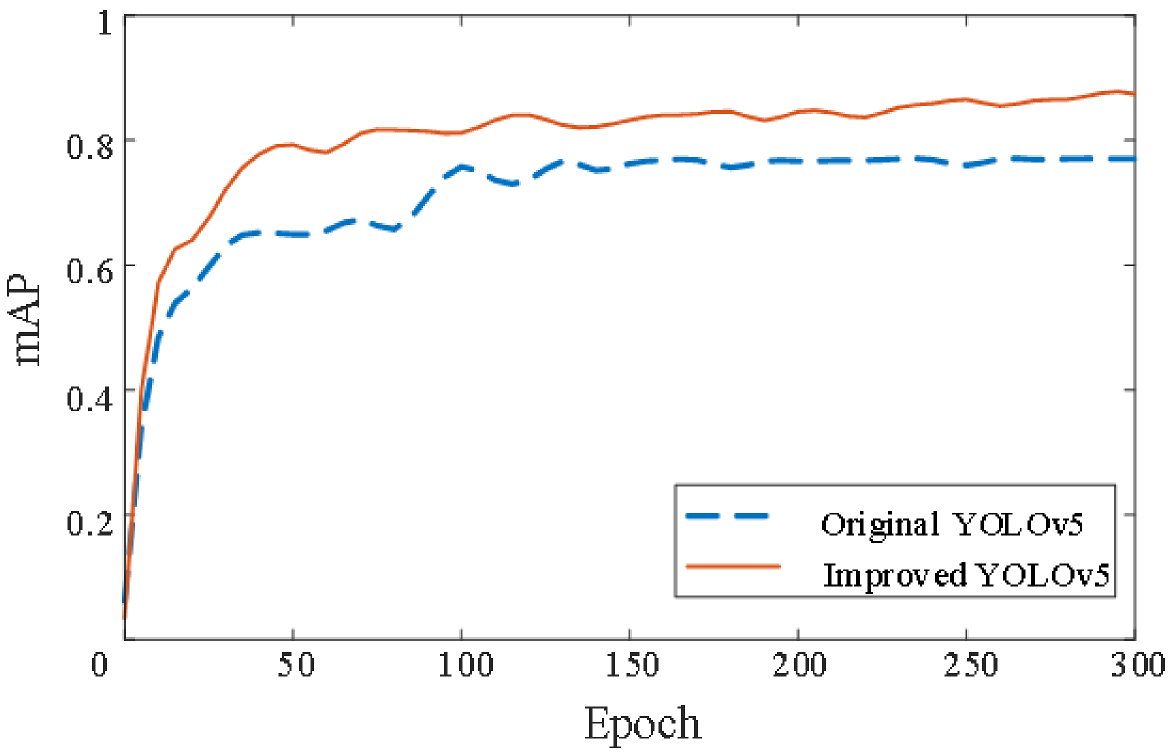


FIGURE 10. Comparison of two YOLOv5s detection models mAP@0.5.

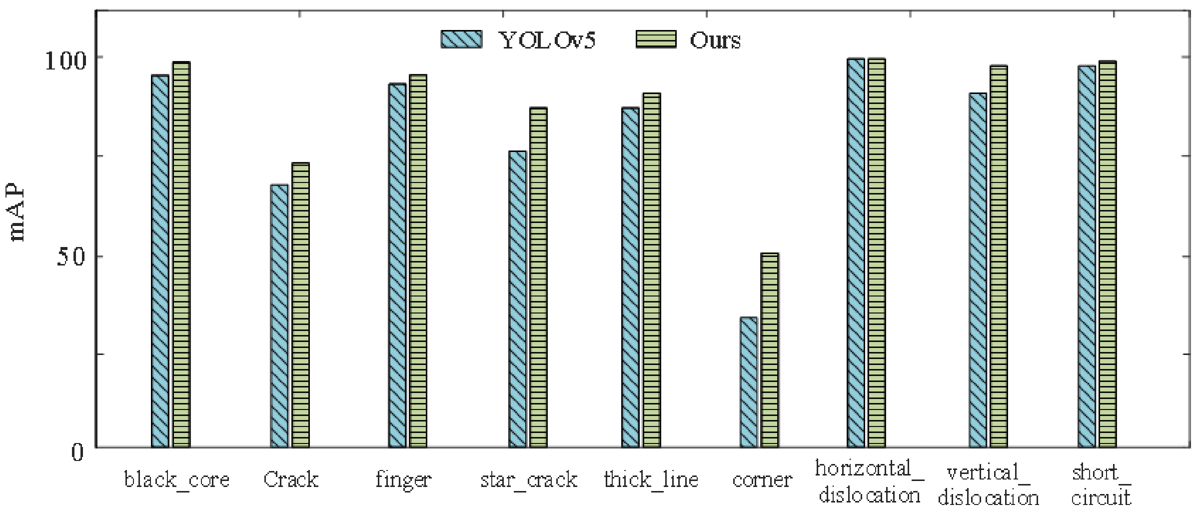


FIGURE 11. Comparison of the average accuracy of the original YOLOv5 and optimized YOLOv5 algorithms for various types of defect detection.

TABLE 3. Detection results of our model on ELPV dataset.

Models	P%	R%	mAP@0.5(%)
ours	98.97	98.23	96.1

detection performance of the optimized YOLOv5 model is higher than that of the original YOLOv5 model.

3) PERFORMANCE COMPARISON EXPERIMENTS

Based on the comparison with the original model, in order to further validate the performance of the improved method proposed in this paper in solar cell surface defect detection, we detected the optimized YOLOv5 detection model together with several mainstream models such as SSD, Faster R-CNN, and BAF-Detector [32] for solar cell surface defects and

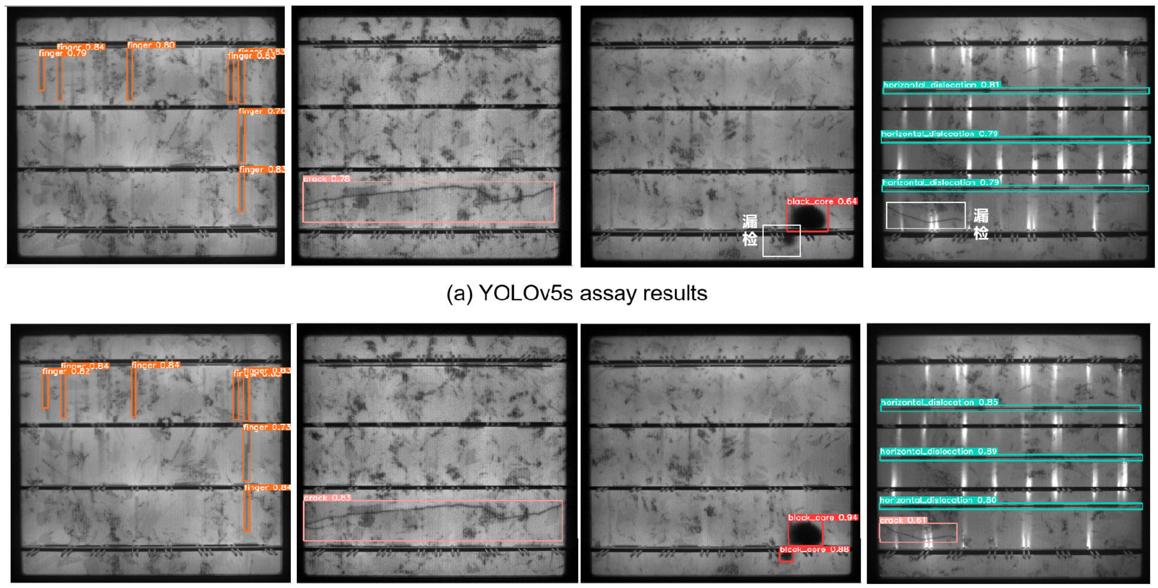


FIGURE 12. Optimized YOLOv5s and the original YOLOv5s detection results comparison.

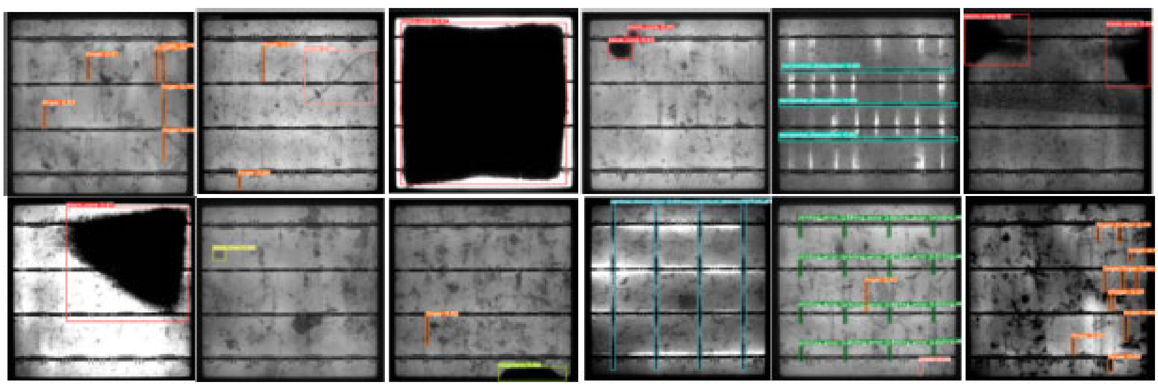


FIGURE 13. Visualization of partial detection results of YOLOv5s after optimization.

performed multivariate analysis on the detection results, and the obtained data are shown in Table 2.

Also to verify the effectiveness of the model, it was trained on the publicly available dichotomous dataset ELPV as well, and the training results are shown in Table 3.

In this paper, P, R and mAP are used as evaluation indexes, and IOU greater than or equal to 0.5 is a positive sample and IOU less than 0.5 is a negative sample in the training process. As can be seen from Table 2, the detection accuracy of the optimized YOLOv5 model is improved by at least 6.63% compared with several other classical models. This is because the models in this paper are optimized with full consideration of the characteristics of the solar cell defect data concentration target, through comparison test, can be concluded that the optimized YOLOv5 model can meet the requirements of solar cell surface defects higher detection accuracy. And the mAP of the model can reach 96.1% when trained on the public dataset, which further verifies the effectiveness of the optimized model. In addition, the optimized model in this paper can accurately detect nine types of defects, and the comparison of YOLOv5 detection accuracy for each type of defects before and after optimization is shown in Figure 11.

From the figure, it can be found that the optimized model has greatly improved the identification and detection of some defect types that are difficult to be detected by other models, including the detection accuracy of star cracks from the original 76% to 87%, and the detection accuracy of black corners from the original 33% to 50%. The detection accuracy of the remaining several defects has also been improved to a greater or lesser extent. In general, the model has a detection accuracy of more than 50% for the nine types of defects in solar cells, and even the detection accuracy of five types of defects such as black core and short circuit has reached about 95%.

Several representative solar cell defect images are randomly selected for detection, and the detection results of the original YOLOv5s detection model and the optimized YOLOv5s model are shown in Figure. Figure 12(a) shows the detection results of the original YOLOv5s model, and Figure 12(b) shows the detection results of the optimized YOLOv5s model. From Figure 12(a), we can see that YOLOv5s has the problems of missed detection and low detection accuracy under the interference of complex background. And in Figure 12(b), the optimized YOLOv5s model in this paper reduces the leakage detection and improves the detection accuracy. The comparison shows that the optimized YOLOv5s model has more accurate detection results, can capture the key information of defects, and has better detection performance.

Finally, a variety of types of solar cell defect images were selected for detection, and the detection results are shown in Figure 13.

V. CONCLUSION

In this paper, an optimized YOLOv5 solar cell surface defect detection model is proposed for solar cell defects that are difficult to collect, difficult to distinguish, easy to mis-detect and miss detection, etc. The model achieves defect detection at different scales by introducing a CA attention mechanism and replacing the decoupling head to enhance the feature extraction capability. Meanwhile, in order to make the model detection ability more effective, this paper adopts a combination of five data enhancement methods, namely Mosaic, Mixup, hsv transform, scale transform and flip, to improve the accuracy of feature training and enhance the robustness of the model. Finally, the comparison experiments and ablation experiments show that the optimized YOLOv5 model not only improves the mAP by 10.38% to 87.4% compared with the original detection model, but also has significant adaptability to accurately detect nine types of defects in solar cells. Meanwhile, in order to further verify the effectiveness of the model, its test mAP reached 96.1% on the public dataset. It indicates that the model has a good application prospect in solar cell defect detection. The direction of future work will be to further optimize the model, further solve the problems of imbalance of defect types in the dataset and difficulty in detecting some defect types, and consider whether it is possible to further improve the detection accuracy and speed of the model by reducing the number of model parameters to make the model more practical.

REFERENCES

[1] S. B. Jha and R. F. Babiceanu, "Deep CNN-based visual defect detection: Survey of current literature," *Comput. Ind.*, vol. 148, Jun. 2023, Art. no. 103911, doi: 10.1016/j.compind.2023.103911.

[2] L. Liu, C. Shen, and A. Van Den Hengel, "Cross-convolutional-layer pooling for image recognition," 2015, *arXiv:1510.00921*.

[3] N. Liu, L. Wan, Y. Zhang, T. Zhou, H. Huo, and T. Fang, "Exploiting convolutional neural networks with deeply local description for remote sensing image classification," *IEEE Access*, vol. 6, pp. 11215–11228, 2018, doi: 10.1109/ACCESS.2018.2798799.

[4] A. Hassan and A. Mahmood, "Convolutional recurrent deep learning model for sentence classification," *IEEE Access*, vol. 6, pp. 13949–13957, 2018, doi: 10.1109/ACCESS.2018.2814818.

[5] X. Ren, Y. Zhou, Z. Huang, J. Sun, X. Yang, and K. Chen, "A novel text structure feature extractor for Chinese scene text detection and recognition," *IEEE Access*, vol. 5, pp. 3193–3204, 2017, doi: 10.1109/ACCESS.2017.2676158.

[6] W. Liu, A. Dragomir, E. Dumitru, S. Christian, R. Scott, C.-Y. Fu, and C. B. Alexander, "SSD: Single shot MultiBox detector," in *Computer Vision—ECCV*. Cham, Switzerland: Springer, 2016, pp. 21–37.

[7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[8] C. V. Dung and L. D. Anh, "Autonomous concrete crack detection using deep fully convolutional neural network," *Autom. Construct.*, vol. 99, pp. 52–58, Mar. 2019, doi: 10.1016/j.autcon.2018.11.028.

[9] G. H. Yao and X. C. Wu, "Halcon-based solar panel crack detection," in *Proc. IEEE WCMEIM* Shanghai, China, Nov. 2019, pp. 733–736.

[10] V. S. Bharath, A. Haque, and M. A. Khan, "Fault classification for photovoltaic modules using thermography and image processing," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting*, Baltimore, MD, USA, Sep. 2019, pp. 1–6.

[11] H. Chen, H. Zhao, D. Han, W. Liu, P. Chen, and K. Liu, "Structure-aware-based crack defect detection for multicrystalline solar cells," *Measurement*, vol. 151, Feb. 2020, Art. no. 107170, doi: 10.1016/j.measurement.2019.107170.

[12] J. Balzategui, L. Eciolaza, and N. Arana-Arexolaleiba, "Defect detection on polycrystalline solar cells using electroluminescence and fully convolutional neural networks," in *Proc. IEEE SII* Honolulu, HI, USA, Jan. 2020, pp. 949–953.

[13] A. V. Zubair, B. Yoann, D. Priya, S. Arcot, T. Thorsten, and H. Ziv, "Localization of defects in solar cells using luminescence images and deep learning," in *Proc. IEEE PVSC*, Fort Lauderdale, FL, USA, Jun. 2021, pp. 745–749.

[14] Y. Wang, L. Li, Y. Sun, J. Xu, Y. Jia, J. Hong, X. Hu, G. Weng, X. Luo, S. Chen, Z. Zhu, J. Chu, and H. Akiyama, "Adaptive automatic solar cell defect detection and classification based on absolute electroluminescence imaging," *Energy*, vol. 229, Aug. 2021, Art. no. 120606, doi: 10.1016/j.energy.2021.120606.

[15] A. S. Al-Waisy, D. A. Ibrahim, D. A. Zebari, S. Hammadi, H. Mohammed, M. A. Mohammed, and R. Damasevicius, "Identifying defective solar cells in electroluminescence images using deep feature representations," *PeerJ Comput. Sci.*, vol. 8, p. e992, May 2022, doi: 10.7717/peerj-cs.992.

[16] M. Zhang and L. Yin, "Solar cell surface defect detection based on improved YOLO v5," *IEEE Access*, vol. 10, pp. 80804–80815, 2022, doi: 10.1109/ACCESS.2022.3195901.

[17] L. Li, Z. Wang, and T. Zhang, "GBH-YOLOv5: Ghost convolution with BottleneckCSP and tiny target prediction head incorporating YOLOv5 for PV panel defect detection," *Electronics*, vol. 12, no. 3, p. 561, Jan. 2023, doi: 10.3390/electronics12030561.

[18] S. Prabhakaran, R. A. Uthra, and J. Preetharoselyn, "Deep learning-based model for defect detection and localization on photovoltaic panels," *Comput. Syst. Sci. Eng.*, vol. 44, no. 3, pp. 2683–2700, 2023, doi: 10.32604/csse.2023.028898.

[19] A. Chen, X. Li, H. Jing, C. Hong, and M. Li, "Anomaly detection algorithm for photovoltaic cells based on lightweight multi-channel spatial attention mechanism," *Energies*, vol. 16, no. 4, p. 1619, Feb. 2023, doi: 10.3390/en16041619.

[20] M. Bie, Y. Liu, G. Li, J. Hong, and J. Li, "Real-time vehicle detection algorithm based on a lightweight you-only-look-once (YOLOv5n-L) approach," *Exp. Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 119108, doi: 10.1016/j.eswa.2022.119108.

[21] T.-Y. Lin, D. Piotr, G. Ross, K. M. He, H. Bharath, and B. Serge, "Feature pyramid networks for object detection," in *Proc. IEEE CVPR*, Honolulu, HI, USA, Jul. 2017, pp. 936–944.

[22] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Honolulu, HI, USA, Jun. 2018, pp. 8759–8768.

[23] D. Zou, Y. Cao, Y. Li, and Q. Gu, "The benefits of mixup for feature learning," 2023, *arXiv:2303.08433*.

[24] G. Song, Y. Liu, and X. Wang, "Revisiting the sibling head in object detector," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 11560–11569.

[25] Y. Wu, Y. Chen, L. Yuan, Z. Liu, L. Wang, H. Li, and Y. Fu, ''Rethinking classification and localization for object detection,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 10183–10192.
[26] J. Hu, L. Shen, and G. Sun, ''Squeeze-and-excitation networks,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
[27] J. Park, S. Woo, J.-Y. Lee, and I. S. Kweon, ''BAM: Bottleneck attention module,'' 2018, *arXiv:1807.06514*.
[28] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, ''CBAM: Convolutional block attention module,'' in *Proc. ECCV*, 2018, pp. 3–19.
[29] Q. Hou, D. Zhou, and J. Feng, ''Coordinate attention for efficient mobile network design,'' 2021, *arXiv:2103.02907*.
[30] C. Buerhop-Lutz, S. Deitsch, A. Maier, F. Gallwitz, S. Berger, B. Doll, J. Hauch, C. Camus, and C. J. Brabec, ''A benchmark for visual identification of defective solar cells in electroluminescence imagery,'' in *Proc. Eur. Photovoltaic Sol. Energy Conf. Exhib.*, 2018, pp. 1287–1289.
[31] B. Su, Z. Zhou, and H. Chen, ''PVEL-AD: A large-scale open-world dataset for photovoltaic cell anomaly detection,'' *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 404–413, Jan. 2023, doi: 10.1109/TII.2022.3162846.
[32] B. Su, H. Chen, and Z. Zhou, ''BAF-detector: An efficient CNN-based detector for photovoltaic cell defect detection,'' *IEEE Trans. Ind. Electron.*, vol. 69, no. 3, pp. 3161–3171, Mar. 2022, doi: 10.1109/TIE.2021.3070507.

KAIXING WU received the bachelor's degree in industrial electrical automation from the China University of Mining and Technology, in 1986, and the master's degree in automation from Tianjin University, in 1996. He has undertaken more than 20 scientific research projects as the host or a principal researcher. He has Published more than 80 articles in domestic and foreign journals and two textbooks as a chief (deputy) editor. His research interests include intelligent information processing and information security, network technology, and information security. He is the Director of the Hebei Automation Society and a member of the China Coal Industry Association.

SHA LU graduated from Dezhou University, Dezhou, China, in 2021. She is currently pursuing the master's degree with the Hebei University of Engineering, Handan, China. Her research interests include deep learning-based image processing and object detection in computer vision.

JINXIN CHEN graduated from Changzhi University, Changzhi, China, in 2021. He is currently pursuing the master's degree with the Hebei University of Engineering, Handan, China. His research interest includes deep learning-based image processing in computer vision.