

PVEL-AD: A Large-Scale Open-World Dataset for Photovoltaic Cell Anomaly Detection

Binyi Su, Zhong Zhou, and Haiyong Chen

Abstract—The anomaly detection in photovoltaic (PV) cell electroluminescence (EL) image is of great significance for the vision-based fault diagnosis. Many researchers are committed to solving this problem, but a large-scale open-world dataset is required to validate their novel ideas. We build a PV EL Anomaly Detection (PVEL-AD^{1,2,3,4}) dataset for polycrystalline solar cell, which contains 36,543 near-infrared images with various internal defects and heterogeneous background. This dataset contains anomaly-free images and anomalous images with 10 different categories. Moreover, 37,380 ground truth bounding boxes are provided for 8 types of defects. We also carry out a comprehensive evaluation of the state-of-the-art object detection methods based on deep learning. The evaluation results on this dataset provide the initial benchmark, which is convenient for follow-up researchers to conduct experimental comparisons. To the best of our knowledge, this is the first public dataset for PV solar cell anomaly detection that provides box-wise ground truth and focuses on industrial application. Furthermore, this dataset can also be used for the evaluation of many computer vision tasks such as few-shot detection, one-class classification and anomaly generation.

Index Terms—photovoltaic cell, anomalous dataset, deep learning, defect detection, real-world application

I. INTRODUCTION

PHOTOVOLTAIC solar cells are the main products that can convert solar energy into electric energy. However, the inevitable defects can notably decrease the photoelectric conversion efficiency and lifespan of the modules [1], and then it will cause great economic losses. This can be avoided by continuous inspection of solar cells in intelligent manufacturing process. Manual inspection of EL images is very cumbersome and requires professional knowledge. Therefore, the vision-based automated inspection [2]–[6] process is more desirable.

However, vision-based inspection algorithms often lack defective samples, which are crucial to train an excellent deep-learning model. In order to satisfy the urgent requirement,

This work was supported by National Key R&D Program of China under Grant 2018YFB2100603, Natural Science Foundation of China under Grant 61872024, 62073117, U21A20482 and Central Government Guides Local Science and Technology Development under Grant 206Z1701G. (Corresponding authors: Zhong Zhou; Haiyong Chen).

B. Su, Z. Zhou are with the State Key Laboratory of Virtual Reality Technology and Systems, School of Computer Science and Engineering, Beihang University, Beijing 100191, China (e-mail: Subinyi@buaa.edu.cn; zz@buaa.edu.cn).

H. Chen is with the School of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin 300130, China (e-mail: haiyong.chen@hebut.edu.cn).

¹IEEEDataPort-DOI:10.21227/pz6t-3s77

²<https://github.com/binyisu/PVEL-AD>

³<http://aihebut.com/col.jsp?id=118>

⁴<http://nave.vr3i.com/download/view.action>

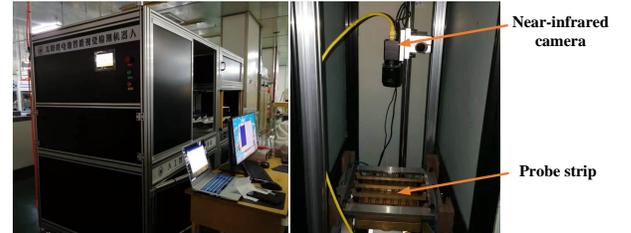


Fig. 1: The left is PV cell intelligent anomaly detection system, right is the near-infrared EL image acquisition subsystem.

we propose a novel comprehensive, large-scale photovoltaic (PV) cell electroluminescence (EL) image anomaly detection dataset, named as PVEL-AD. To build this dataset, EL imaging technology [7] plays a key role in the process of image acquisition. Because, some anomalies are often in the interior of photovoltaic cells, which cannot be photographed directly by optical or infrared camera. But visual inspection using EL imaging technology allows to easily identify anomaly inflicted to solar cells either by external environmental influences such as bump during the manufacturing process, or due to the prior material defects.

EL imaging requires a special environment, which is illustrated in Fig. 1. As we can see, the left is the PV cell intelligent defect detection system, which is designed for the real-world industrial application. The right is the EL image acquisition subsystem, the detailed internal structure is shown in the Fig. 2. In dark room, when PV cell is transmitted to the bottom of probe strip in a suitable position, the sensor will send signal to programmable logic controller (PLC). Then, PLC controls the probe strip to descend until it slightly touches the PV cell. The contact force should not be too large to prevent the solar cell from being damaged. Next, the solar cell that connects to the positive pole is powered on 24-V direct-current (dc) voltage and 8-A current. Excited with the voltage, the solar cell emits near-infrared light in a wavelength around 1000-1200 nm [7], which will be captured by a cooled silicon charge-coupled device (Si-CCD) camera. As illustrated in Table I, the cooled Si-CCD camera is composed of a near-infrared Mono Chrome camera of WP-US146 with a SONY ICX825 chip and an industrial lens of VTG1214-M4. The resolution is about 1 million (1024×1024), which can ensure that the clear image can be captured by the camera. Moreover, due to the weak near-infrared light emitted by the photovoltaic solar cell, the exposure time (0.7 s) should be relatively long. After being captured, EL image will be saved to the computer server for dataset construction. Simultaneously, the probe strip rises

TABLE I: Parameters of the Camera.

Parameter	Value
Type	WP-US146
Chip	SONY ICX825
Resolution	1024×1024
Industrial lens	VTG1214-M4
Exposure	0.7 s
Distance to solar cell	350 mm
Field of View	158×158 mm
Color	Mono Chrome camera

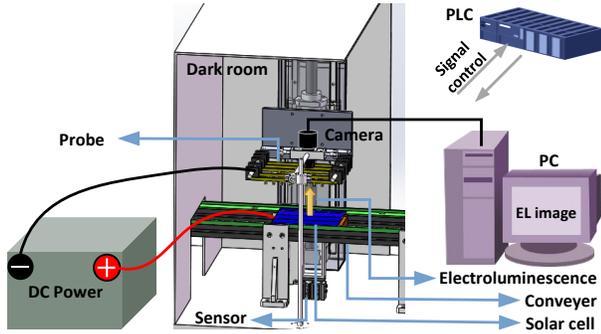


Fig. 2: The internal structure of EL imaging subsystem.

and another solar cell is transmitted by conveyor to repeat above process. Each image is about 1 s from acquisition to storage, which can satisfy the real-time requirements of industrial production.

Several samples of the solar cell EL images are visualized in Fig. 3. The anomalies contain ten types such as linear crack, star crack, finger interruption, black core, misalignment, thick line, scratch, fragment, corner, and material defects. The data in EL imaging that includes a large number of samples and a variety of anomaly types is not cheap. Acquisition equipment manufacturing and the scarcity of defects are hindering factors, which cause a hindrance to the development and evaluation of visual inspection algorithms. Moreover, most researchers work with their own datasets that are not public. The comparison between different inspection approaches is therefore practically not possible.

To support good scientific practices and promote the development of solar cell anomaly inspection approaches, we build a comprehensive and large-scale solar cell dataset (PVEL-AD), which consists of box-annotated EL images of solar cells. All images are labeled with the help of experts, who annotate the position and category of the appearing anomalies. PVEL-AD dataset can be used to evaluate many new proposed ideas and promote the development of solar cell quality monitoring, which motivate us to release the large-scale and multi-category dataset. Using the provided data, we build an anomaly-detection benchmark, which can be convenient for the follow-up researchers to conduct experimental evaluation. The main contributions of this paper are summarized as follows.

1) We build a large-scale open-world dataset (PVEL-AD) for PV cell defect inspection in near-infrared image data. It is collected from the practical industrial manufacturing and consists of 36,543 high-resolution images of one defect-free type and ten defective categories. We provide box-

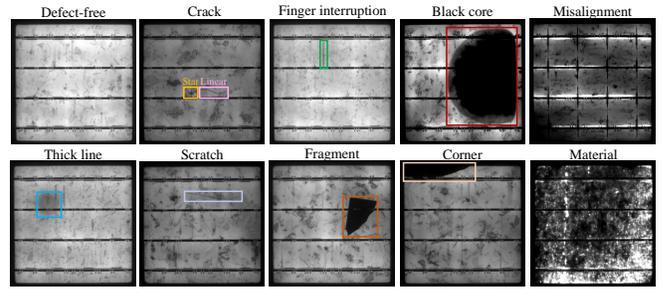


Fig. 3: Example images of PVEL-AD dataset. One defect-free image and nine images that contain ten types of anomalies are presented.

wise ground truths for 37,380 anomalies that can be used to evaluate the newly proposed anomaly detection methods. Additionally, this dataset is also suitable for evaluation of one-class classification, few-shot detection, and defect generation.

- 2) We propose a defect anomaly annotation method, called fine-tuning labelImg (F-labelImg), which can reduce the workload of labeling and prevent some anomalies from being missed.
- 3) We conduct a comprehensive assessment of current state-of-the-art methods for PV cell defect detection task. Furthermore, we show that there are many works that have not been done in this dataset, it still has much room for improvement.

This paper is organized as follows: Section II presents an overview of the related works for existing datasets and detection methods. Section III gives the description of the dataset and application analysis. Section IV shows extensive experiments for anomaly detection. Finally, Section V concludes this paper.

II. RELATED WORKS

In this section, we firstly introduce the datasets that are commonly utilized for anomaly inspection in various scenes and demonstrate the necessity of our dataset. Then, we give a brief overview for the CNN-based anomaly detection methods.

A. Anomaly Datasets

The number of dataset used for the evaluation of anomaly detection methods is still very small. These datasets have few defect categories and defect data. As far as we know, there still not exist a comprehensive, large-scale, and number-increasing solar-cell dataset for open-world practical scene, such as our PVEL-AD dataset.

Song *et al.* [8] provided a steel surface defect dataset, which includes 1,800 gray-scale images with six types of defects. Each class of defect has 300 samples with a resolution of 200×200 pixels. This dataset only provides defective samples, while the non-defective samples are ignored. Li *et al.* [9] proposed a CrackForest Dataset. There are 2,688 images of bridge cracks without any annotation. It needs to be further processed and cannot be used to evaluate the new proposed

TABLE II: Statistical Comparison with Different Datasets.

	Dataset	Anomaly image	Anomaly-free image	Total	Resolution	Anomaly Types	Box or pixel label
Song <i>et al.</i> [8]	steel surface defect dataset	1800	-	1800	200×200	6	yes
Li <i>et al.</i> [9]	CrackForest dataset	2688	-	2688	1024×1024	1	no
Carrera <i>et al.</i> [10]	NanoTWICE dataset	40	5	45	1024×696	1	yes
Tabernik <i>et al.</i> [11]	Kolektor Surface Defect dataset	52	347	399	704×256	1	yes
Bergmann <i>et al.</i> [12]	MVTec AD dataset	1258	4096	5354	700~1024	73	yes
Buerhop <i>et al.</i> [13]	elpv dataset	918	1508	2426	300×300	3	no
Ours	PVEL-AD dataset	25192	11351	36543	1024×1024	10	yes

algorithms directly. Carrera *et al.* [10] introduced NanoTWICE dataset, which only consists of a single-class defect with 40 samples and 5 anomaly-free samples. The single class and small number of data limit the generalization evaluation of the new proposed algorithms. For few-shot defect detection, Tabernik *et al.* [11] introduced a Kolektor Surface-Defect dataset. It was collected in a controlled industrial environment under the open-world situation. This dataset consists of 399 samples (52 anomalous samples and 347 anomaly-free samples). Due to the small number of defective samples, this dataset is suitable for the few-shot defect detection task.

In recent years, Bergmann *et al.* [12] introduced a comprehensive real-world dataset (MVTec AD dataset) for unsupervised anomaly inspection. MVTec AD dataset contains 5,354 high-resolution colorful images of different object and texture categories. Although this dataset has multi-category samples, it is artificially purposeful mirroring industrial data, rather than the real industrial dataset, which may contain many unexpected anomalous samples. For the identification of PV cell defect, Buerhop *et al.* [13] proposed a EL image dataset (elpv dataset), which included 2,426 samples with class annotations. These samples were cropped from 44 different raw images, of which 26 are polycrystalline and 18 are monocrystalline. Due to the small number of raw images in elpv dataset, the diversity of the defects is limited.

A challenging problem for industrial defect inspection is that the defect samples are not easy to be collected. The actual dataset collection process is expensive and takes a long time. Several workers spent 2 years collecting our PVEL-AD dataset, which consists of 11,351 anomaly-free samples, 21,044 defective samples with box-wise annotations, and 4,148 defective samples with category annotations. A comparison between different datasets is carried out in Table II. It is not hard to see that our PVEL-AD dataset has much more anomaly and anomaly-free images than other industrial datasets [8]–[13]. Moreover, compared with the same type dataset (elpv), PVEL-AD dataset has higher resolution, and the anomaly types are more diverse and comprehensive.

B. Anomaly Detection

In last three years, many methods have been proposed to detect the PV cell anomaly. We restrict a brief overview to current related works for PV cell anomaly detection.

Since Buerhop *et al.* [13] proposed a benchmark based on elpv dataset for optical identification of defective PV cells in EL images, many approaches were proposed based on this dataset. Deutsch *et al.* [14] firstly tried to use an

TABLE III: Statistical Overview of the PVEL-AD-2019 Dataset.

Vision task	Category	Anomaly number			Image number
		Train	Test	Total	
Anomaly detection	Crack	452	685	1137	2129
	Finger	592	1249	1841	
	Black core	251	272	523	
Classification	Anomaly-free	-	-	-	1500

TABLE IV: Statistical Overview of the PVEL-AD-2021 Dataset.

Vision task	Category	Anomaly number			Image number
		Train	Test	Total	
Anomaly detection	Linear crack	1260	2797	4057	21044
	Finger	2957	22636	25593	
	Black core	1028	3877	4905	
	Thick line	981	1585	2566	
Few-shot anomaly detection	Star crack	135	83	218	901
	Corner	9	12	21	
	Fragment	7	5	12	
	Scratch	5	3	8	
one-class Classification and anomaly generation	Misalignment	-	-	-	3247
	Material	-	-	-	11351
one-class Classification and anomaly generation	Anomaly-free	-	-	-	11351

end-to-end deep learning network to accomplish the two-class classification task in elpv dataset, which achieved a more accurate result than traditional methods. Akram *et al.* [15] proposed a light-weight CNN model to classify defect or defect-free images in elpv dataset, which achieves better performance than [14]. Qian *et al.* [16] combined short term and long term feature extracted by pre-trained CNN to detect micro-cracks in elpv dataset. This method can effectively detect various kinds of micro-cracks. Ge *et al.* [17] designed a Hybrid Fuzzy Convolutional Neural Network (HFCNN) to recognize the defective PV cell image, which achieved the state-of-the-art results in elpv dataset. However, elpv dataset does not provide box-wise or pixel-wise annotations that play a vital role in fair comparison between new developed anomaly inspection methods. Moreover, as illustrated in Table II, the number and types are relatively smaller than our PVEL-AD dataset. Thus, it is hard to train a high-efficiency model.

Some scholars developed several methods on their own PV cell dataset. Liu *et al.* [3] introduced a novel iterating tensor voting method to segment the crack defect in 405 EL

images. Chen *et al.* [5] proposed a novel steerable evidence filter (SEF) to detect micro-crack in PV cell EL image, where 572 images were used to evaluate this method. To detect the surface defect in PV cell, Chen *et al.* [18] designed a multi-spectral deep CNN model to inspect defect with six categories. Their dataset include 21,245 images that is relatively large. Subsequently, Chen *et al.* [19] developed an end-to-end CNN architecture with attention mechanism to locate the PV cell defect with six categories. Du *et al.* [20] introduced several CNN models to classify PV cell in their dataset with 1440 thermal images. Recently, some researchers proposed several pixel-wise solar cell defect inspection methods. Han *et al.* [21] proposed a two-stage defect segmentation method for polycrystalline solar cell based on deep convolution network. This approach achieved a good performance for micro-crack inspection. Rahman *et al.* [22] proposed a novel modified U-net architecture to segment various complex defects in EL images. This method outperformed many previous methods under complex background disturbance.

All the above works are done on small-scale elpv or their own datasets. There is an urgent need for a large-scale and public dataset to fairly evaluate the newly proposed defect detection methods, such as our PVEL-AD dataset.

For our large-scale PVEL-AD dataset, we have done several works based on part of it, but there is still considerable room for improvement. Su *et al.* [2] proposed a bag-of-feature model to classify defects in PV EL image. This approach achieved a high accuracy for image classification, but it was used to classify the low-resolution cropped images (128×128 pixels). To detect defect in raw EL image (1024×1024 pixels), Su *et al.* [4] proposed a novel region-based deep learning method (Faster RPN-CNN), which employed an attention-based region proposal network (RPN) to extract the suspected defect region in raw EL image, and a following fully connected network was applied to output the specific defect category and position. However, this approach cannot solve gradient vanishing problem of the small scale defect as the network deepens. To solve this problem, Su *et al.* [6] developed a bidirectional attention-based feature pyramid network (BAFPN) to accomplish multi-scale feature fusion. BAFPN improved the robustness of network to scales, thus the proposed detector achieved a good performance in multi-scale defect detection task.

III. DATASET DESCRIPTION AND APPLICATION

In this section, we sequentially describe dataset, annotation tool, and dataset application.

A. Dataset

A comprehensive and high-quality dataset often has a great value in promoting the development of an industry such as PV cell manufacturing. Curiously, there is an absence of comprehensive real-world dataset available for such scenarios. The proposed PVEL-AD dataset is established in the process of practical intelligent manufacturing. It is an open-world dataset, that's mean as time goes by, data is continuously accumulated, and new anomalies outside the original dataset

may appear. Several workers have spent two years to collect it. This dataset is unique in its large number and variety of PV-cell samples, as presented in Fig. 3 and Table II. In contrast to the elpv dataset that was proposed by [13], PVEL-AD dataset has more categories and data. Moreover, 8 types of defects are provided box-wise ground truth that labels the accurate defect category and position in the image.

1) *PVEL-AD-2019*: PVEL-AD-2019 is collected and annotated by 2019. As illustrated in Table III and Table IV, this dataset is a small dataset compared with the released PVEL-AD-2021 in this paper. The total image number is 3,629, which is about one tenth of the proposed PVEL-AD-2021. Although some works [4], [6] has been done in the PVEL-AD-2019, but note that this dataset is not released in public until now.

2) *The proposed dataset PVEL-AD-2021*: As time goes by, image samples are continuously collected and accumulated in the process of real-world industrial manufacturing. The proposed PVEL-AD-2021 dataset is incremented based on the original PVEL-AD-2019 dataset. The number of annotated images is increased from 2,129 to 21,044. The anomaly-free image is increased from 1500 to 11351. Moreover, new categories are introduced in PVEL-AD-2021 dataset, such as thick line, corner, fragment, scratch, misalignment, and material anomalies. Note that if there is no special illustration in this paper, PVEL-AD stands for PVEL-AD-2021.

The image in PVEL-AD dataset has some special characteristics such as near-infrared image, heterogeneous background. The material of solar cell is polycrystalline silicon, which shows the advantages of low cost and high efficiency. For this material, the EL imaging technology is a convenient and fast internal defect visualization method. PVEL-AD dataset is collected by EL imaging technology in the practical industrial manufacturing. It consists of 36,543 high-resolution images (1024×1024 pixels) that can be divided into one defect-free type and ten defective categories. The anomalies with box-wise annotations consist of eight types (linear crack, star crack, finger interruption, black core, thick line, scratch, fragment, and corner) with box-wise annotations. Linear crack, finger interruption, black core, and thick line are high-frequency defects that can be used to evaluate the conventional anomaly detection methods. The training and testing dataset distribution for anomaly detection is presented in Fig. 4. The testing samples are slightly larger than the training samples. The reason why we divide the dataset like this is that the training data has covered rich texture and shape, which are adequate to train a high-efficiency deep learning model. Moreover, we also provide a large number of defect-free images, the follow-up researchers can use defect-free images to augment defective images through GAN-based methods [27], [28] or other augmentation strategies such as CutMix [29], which will further improve the accuracy of defect detection.

There are four few-shot anomalies: star crack, scratch, fragment and corner. Compared with other anomalies, these defects rarely appear in the manufacturing process. They are promising to evaluate the few-shot anomaly detection approaches [23]. The remaining two types of anomalies (Misalignment and material) are not suitable for object detection, but are suitable for image classification. Additionally, 11,351

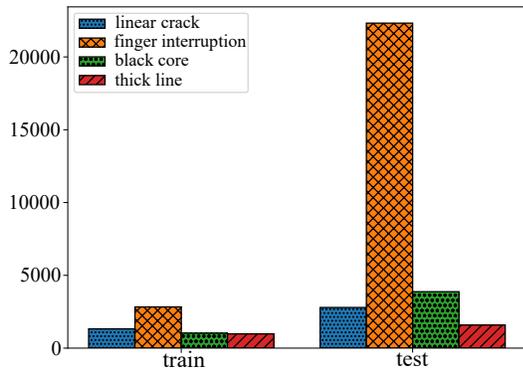


Fig. 4: The data distribution of anomaly detection for training and testing.

anomaly-free images are also collected into this dataset, they are one of the characteristics for anomaly detection dataset, which is different with natural scene dataset such as PASCAL VOC [25] in which all the images contain object. These anomaly-free images that do not contain anomalies are very valuable for one-class-classification [26] or anomaly generation [27].

B. Annotation Tool

The annotation tool is used to annotate the category and position of anomalous defects in the image, it provides the ground truth to train the deep learning model. To reduce the workload of labeling and prevent some background interference anomalies from being missed, we develop a simple defect anomaly annotation method for large-scale dataset, called fine-tuning labelImg (F-labelImg). The flowchart of F-labelImg is shown in Fig. 5. In this paper, the pre-trained CSPDarkNet-s-m-l (small, middle, large) [30] are employed as the feature extraction network of anomaly detector (YOLOv5 [32]). With the increase of labeled data, the feature extraction network gradually deepens from CSPDarkNet-s to -m to -l. The parameter setting and data augmentation for training are same as Section IV. A. Specially, we firstly employ the total original PVEL-AD-2019 dataset to train the YOLOv5-s detector, which is used as initialization model for annotation. The initialized detector is used to predicted 5000 images for unlabeled data, and save the detection results to text file for each image. The category and position in the text files are encoded into xml files (PASCAL VOC format [25]) for each image. Next, the labelImg tool maps the detected results into each image. Then, we can get the ground truth by manually fine-tuning the boundary and category of the box or deleting false detected box or adding missing detected box. Please note that not only the number of annotation files is increasing during the annotation process, but also the types of defects. Until we get 5000 new annotated images, we use all annotated images to train YOLOv5-m, the remaining images are used for prediction, and repeat the above process. Until another 5000 new annotated images are obtained, we use all annotated images to train yolov5-l, and the remaining images are used for prediction. Finally, we annotate the rest of dataset. After

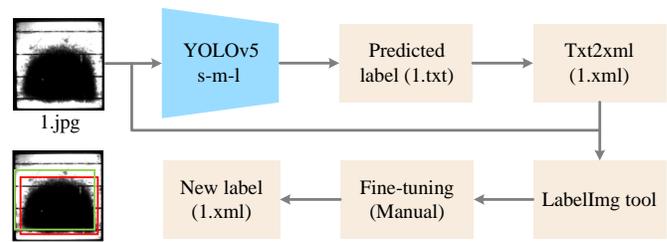


Fig. 5: The flowchart of F-labelImg. Green box is predicted by model, red box is fine-tuned by ourselves.

TABLE V: Similarity Between the Image Labeled by Original labelImg and F-labelImg. P+F Denotes the Prediction and Fine-tune.

Similarity	labelImg	F-labelImg (P+F)
mIoU		99.13%
Time (h)	1.5	0.5

repeated screening and elimination of PVEL-AD dataset, high-quality dataset is finally obtained, which is convenient to carry out research works. The difference between F-labelImg and traditional labelImg is that F-labelImg is a semi-automatic labeling method, which uses a pretrained model to obtain a posterior bounding box in advance, and then performs manual fine-tuning.

To validate the effectiveness of the F-labelImg, a simple method is used to compare the similarity between the image labeled by F-labelImg and the image labeled manually (labelImg). The similarity can be evaluated by mean Intersection over Union (mIoU) index. With the assistant of YOLOv5-s trained by original PVEL-AD-2019, we label 1000 defective images that includes 1536 anomalies to evaluate the proposed F-labelImg. As illustrated in Table V, the mIoU between the image labeled by F-labelImg and the image labeled manually is 99.13%, which presents that there is almost no difference between the image labeled by F-labelImg and the image labeled manually. The proposed F-labelImg can acquire high-quality ground truth. Moreover, F-labelImg is three times faster than labelImg. The reason is that for the evaluation of F-labelImg tool, the mode that used to infer the images is YOLOv5-s, which takes about nine seconds to predict 1000 images. However, the fine-tuning of the predicted boxes speeds half an hour. Thus, inference time plus fine-tuning time is the annotation time for F-labelImg tool, which is about half an hour. For labelImg tool without the help of YOLOv5-s, we would take a lot of time to search for defects with naked eyes, especially those small defects with serious background interference. Thus, it speeds more time (an hour and a half). This reveals that F-labelImg is more efficient than labelImg to annotate the solar cell EL images. The more accurate the model labeling, the less our workload will be, but it would hardly affect the final labeling accuracy.

C. Dataset Application

1) **Anomaly detection:** as mentioned above, we provide box-wise annotations for eight types of anomalies. As is

illustrated in Table IV, due to the unbalanced or long-tail distribution of the labeled data, only linear crack, finger interruption, black core and thick line are suitable to training a traditional anomaly detection model.

- 2) **Few-shot object detection:** to solve the few-shot object detection problem, many few-shot object detection (FSOD) methods [23], [24] have been proposed. In FSOD, there are base classes in which sufficient objects are annotated with bounding boxes. But very few labeled objects are available in novel classes. The novel class set does not share common classes with the base class set. With the aid of abundant data in base classes, the few-shot detectors are expected to learn from limited data in novel classes and detect all novel objects in a held-out testing set.
- 3) **One-class classification:** it is unclear what kinds of defect may appear during the practical manufacturing process. Thus, one-class classification [26] is very suitable for anomaly elimination. The one-class classification model is trained only on a single class of the dataset. As for PVEL-AD dataset, the anomaly-free images are provided to train the one-class classification model. While testing the images, the other classes will be detected as anomalies except for the anomaly-free images.
- 4) **Anomaly generation:** the defective images often rarely appear in the manufacturing process. However, the anomaly-free images are sufficient. Thus, anomaly generation is an more and more popular solution to augment the anomaly data, which employs anomaly-free image to generate realistic anomalous image based on generative adversarial network [27], [28].

IV. BENCHMARK

We limit our research on the anomaly detection and conduct a thorough evaluation of multiple object detection methods [4], [6], [31], [32] as initial benchmark on our PVEL-AD dataset. The benchmark¹ presents the experimental results of the latest methods on PVEL-AD dataset, which aims to prepare a reference for future anomaly detection methods. It can facilitate subsequent researchers to compare their methods with those in the benchmark. Simultaneously, the quality of our dataset can be initially verified by these latest methods.

A. Experimental Setup

1) *Transfer Learning and Data Augmentation:* Transfer learning [32] and data augmentation [31] are the basic operations to train a high-efficiency object detection model. In this paper, transfer learning is used in every anomaly detection method, which extracts texture and semantic feature through the ImageNet-pretrained network. This operation can accelerate the speed of convergence and improve accuracy of the network.

Deep learning is a data-driven approach, which needs a large number of data to train an excellent model. Thus, data augmentation is necessary. It can prevent the model from over-fitting and improve the performance of the model. Data augmentation

is divided before-training stage and during-training stage. In this paper, we conduct horizontal flipping before training to augment the training dataset for above mentioned anomaly detection methods. Thus, the augmented training dataset is twice the size of the original one. Moreover, random data augmentation (resize, place, crop, and distort) is also applied during training process, it can improve the robustness of the model to corresponding transformations.

2) *Anomaly Detection Method:* Several methods [4], [6], [31], [32] proposed recently are applied to carry out a comparison in the proposed PVEL-AD dataset. The specific implementation details for each method are as follows: Faster RPN-CNN [4] is a two-stage anomaly detection network, which suppresses the redundant anomaly-free proposals by introducing the attention mechanism in region proposal network (RPN). The ImageNet-pretrained VGG16 is selected as the feature extraction network. The batch size is set to 1, and the input image is resized to 600×600. The max iteration is fixed to 40,000 during training. Top 50 proposals (score rank) are selected as the input of the following detection network, which will predict the final category and position of the anomalies.

YOLOv5 [32] divides the object detection network into four parts: input, backbone, neck, and head. Fast and accurate detector can be obtained by optimizing each part of object detection network. YOLOv5 integrates a lot of tricks for each part of the network during training and testing. For training, mosaic data augmentation, adaptive anchors, Complete IoU, focus structure, hard swish activation function, path aggregation network, cross stage partial network and so on are conducted to optimize the speed and accuracy of object detection. The batch size is set as 8. Moreover, 300 epochs are used to train the detection model. For testing, test time augmentation (TTA) is applied to improve the robustness and accuracy of the detector. TTA augments each test image by the horizontal flipping and cropping to obtain the augmented images, which are fed to the prediction model and generate multiple predictions. TTA integrates these results to output the final predictions for the test image.

EfficientDet [31] is an one-stage multi-scale network, which views the object detection as a regression task. For this one-to-end detector, the pre-trained efficientnet b0, b1, b2, b3 [33] are employed as the feature extraction network. This detector is not a fixed model. From b0 to b3, the width, depth, and input resolution of the network are gradually raising, and the feature representation ability is enhanced simultaneously. Moreover, the batch size and epoch are set to 8 and 300 respectively. Early stopping and focal loss are also adopted in the training process.

BAF-Detector [6] is a multi-scale anomaly detection network, which predicts the anomalies in every pyramidal layers. The resnet101 [34] is employed to extract the multi-scale features from the input image. Based on these features, the suspected anomalous boxes are predicted by the following Bi-directional Attention Feature Pyramid Network and Region Proposal Network (BAFPN-RPN). Subsequently, a new branch is used to give the final class and position of the anomalies. In BAFPN-RPN, a novel cosine non-local attention module is used to suppress the complex background feature, and

¹Competition benchmark: <https://www.kaggle.com/competitions/pvelad>

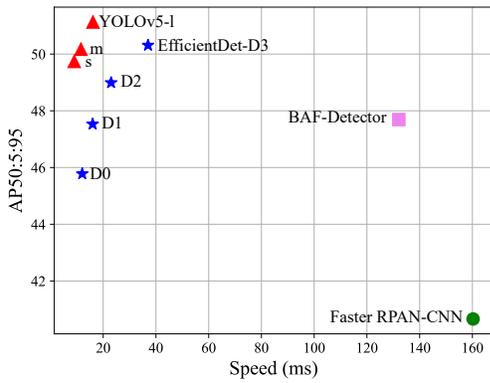


Fig. 6: The performance comparison on PVEL-AD dataset of different methods.

highlight the anomaly feature simultaneously. All input images are resized to 600×600 pixels, and the max iteration is fixed to 40,000 during training. The multi-scale prediction improves the robustness of network to scales, thus BAF-Detector is good at dealing with multi-scale defect detection task.

3) *Evaluation Metric*: Average Precision (AP) and mean Average Precision (mAP) are applied to assess defect detection results. Precision, recall and Intersection over Union (IoU) are used to compute the AP value. The AP is evaluated with different IoUs [6]. It can be calculated for 10 IoUs varying in a range of 50% to 95% with steps of 5%, usually reported as AP50:5:95. It can also be evaluated with single value of IoU, where the most common values are 50% and 75%, reported as AP50 and AP75 respectively. AP is used to evaluate single-class object, mAP is the mean AP value of all classes. Precision/Recall (P/R) curve is an intuitive presentation, which can directly evaluate the effect of the anomaly detection. The AP value is obtained by calculating the area enclosed by the P/R curve and the axes. Parameters number and frames per second (FPS) are the metrics used to assess the time efficiency. All the experiments are carried out in server with a 24G RTX 3090 GPU.

B. Experimental Results and Analysis

Experimental results for anomaly detection are presented in Table VI. The evaluations provide a reference benchmark and validate the quality of the PVEL-AD dataset. Strengths and weaknesses of each approach are discussed on each category of anomaly. It shows that there is still considerable room for improvement of PV EL anomaly detection.

Looking at the experimental results as a whole, no algorithm wins with absolute advantage. In term of mAP of each method, the YOLOv5-l achieves the best anomaly detection performance 51.13% mAP50:5:95 with a relative small parameter number (50.3M). For the speed, the FPS of YOLOv5-s is 111.11 frames per second, which is faster than other methods. Although efficientDet-D0 has the least number of parameter (3.9M), the FPS (83.33) is not the fastest. There are two reasons, one is the repeated use of feature parameters in the feature fusion stage of efficientDet-D0, which consumes computational time. Another is acceleration

TABLE VI: Performance Comparison on PVEL-AD Dataset.

Method	mAP50:5:95	mAP50	mAP75	params	FPS
Faster RPAN-CNN [4]	40.66%	73.24%	32.30%	260.50M	6.24
BAF-Detector [6]	47.69%	80.77%	41.80%	120.87M	7.57
EfficientDet-D0 [31]	45.78%	76.37%	38.40%	3.9M	83.33
EfficientDet-D1 [31]	47.53%	79.88%	42.06%	6.6M	62.50
EfficientDet-D2 [31]	48.99%	80.17%	44.47%	8.1M	43.48
EfficientDet-D3 [31]	50.31%	81.24%	46.67%	12M	27.02
YOLOv5-s [32]	49.74%	81.19%	46.04%	7.10M	111.11
YOLOv5-m [32]	50.17%	81.57%	46.45%	22.0M	86.49
YOLOv5-l [32]	51.13%	81.93%	46.86%	50.3M	62.03

TABLE VII: AP50 of Four Anomalies.

Method	AP50			
	Linear crack	black core	Finger interruption	Thick line
Faster RPAN-CNN [4]	45.52%	92.73%	92.96%	61.77%
BAF-Detector [6]	63.95%	98.25%	94.24%	66.63%
EfficientDet-D3 [31]	64.18%	97.64%	90.42%	72.71%
YOLOv5-l [32]	66.04%	97.84%	93.40%	70.46%

of the focus module [32] adopted by YOLOv5-s, which is designed for parameter reduction and speed increase. By above analysis, we can see that YOLOv5(s-m-l) is a light-weight and high-efficiency object detector. It manage to continuously obtain outstanding performance in practical industrial applications. Fig. 6 shows the performance comparison on PVEL-AD dataset of different methods. In addition to providing a benchmark, the experimental results of these algorithms can also be used to verify the quality of the annotated dataset. As shown in Fig. 6, for EfficientDet, the anomaly detection results are steadily improving from D0 to D3. Simultaneously, the anomaly detection results are also steadily improving from YOLOv5-s to YOLOv5-l. This validates that the annotation quality of the dataset is high. It enables the algorithm to obtain stable experimental results.

The detection result of each anomaly is presented in Table VII. YOLOv5-l achieves the best detection results 66.04% AP50 for linear crack anomaly with a real-time speed (62.03 FPS). Linear crack defect mostly locates on the edge of the PV cell, some of them present small scale under complex background disturbance, which causes its relatively poor detection results than other types of anomalies. BAF-Detector is better at detecting black core and finger interruption anomalies (98.25% and 94.24% AP50 respectively). For thick line anomaly, EfficientDet-D3 achieves the best detection performance 72.71% AP50. Totally, no algorithm wins with absolute advantage in terms of each anomaly.

P/R curve directly presents the detection result for each single category. The area enclosed by P/R curve and axis denotes the AP50 value. The bigger the area is, the larger the AP50 is. As shown in Fig. 7, the first, second, third, and fourth rows are the P/R curves of Faster RPAN-CNN, BAF-Detector, EfficientDet-D3, and YOLOv5-l respectively. The best result is second column (black core), we hope the curve is as close as possible to $y = 1$, which represents that the score-rank anomaly box in the raw image are all predicted correctly. The worst detection result is linear crack anomaly, as is shown in the first column of Fig. 7, the P/R curve decreases faster than

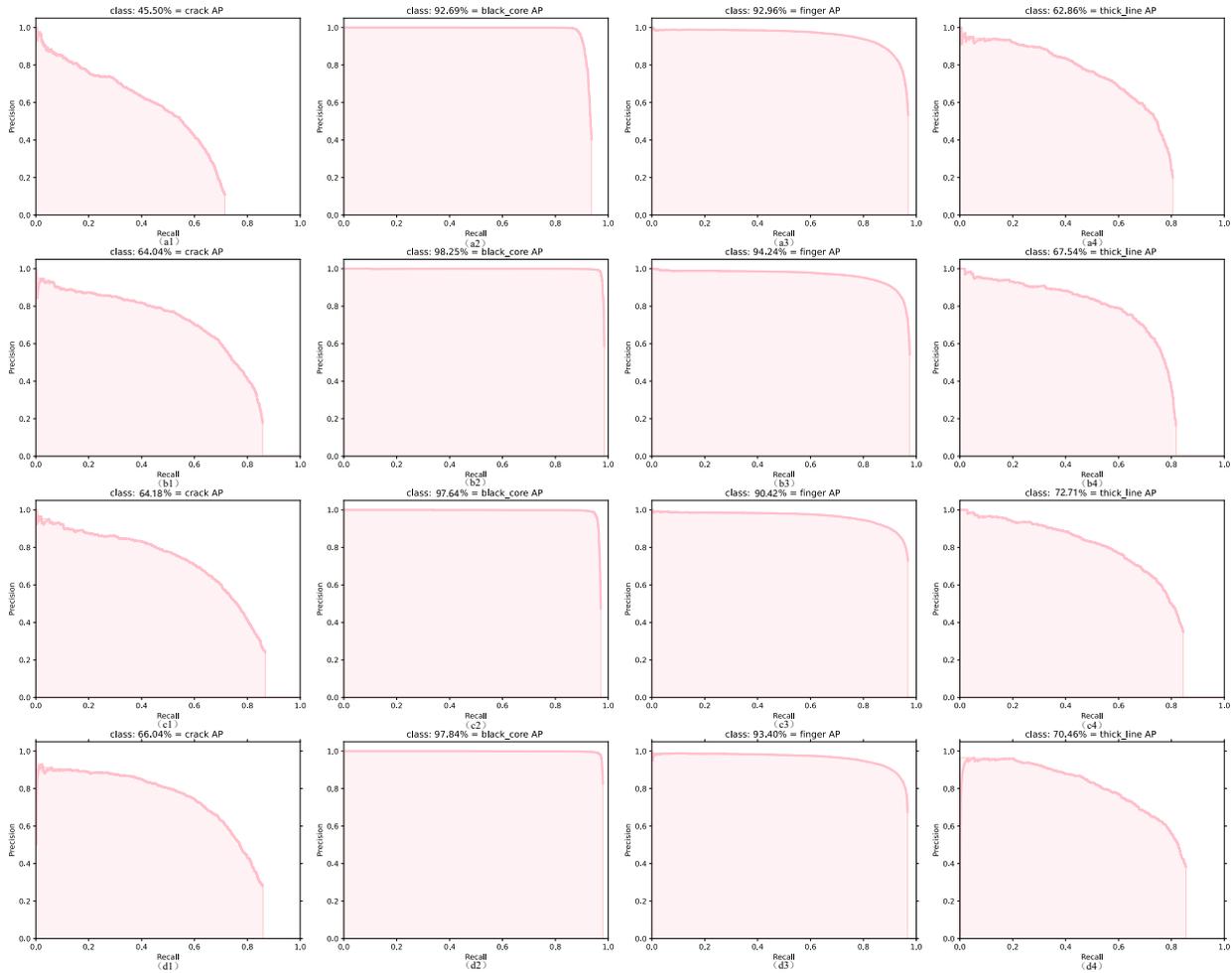


Fig. 7: The Precision/Recall (P/R) curves of different methods. The first, second, third, and fourth rows are the P/R curves of different methods. (a1)-(a4) Faster RPAN-CNN, (b1)-(b4) BAF-Detector, (c1)-(c4) EfficientDet-D3, (d1)-(d4) YOLOv5l.

other types of anomaly. An important reason is that PVEL-AD dataset includes substantive small scale linear cracks, and most of them locate in the edge of the PV cell, which bring some challenges to linear crack anomaly detection. Moreover, (d1), (b2), (b3), and (c4) of Fig. 7 present the best detection results for linear crack, black core, finger interruption, and thick line respectively. As analyzed above, in our proposed PVEL-AD dataset, black core and finger interruption are easy to be detected. Linear crack and thick line are challenging anomaly, which are hard to be detected by CNN-based detector. There is still a large room for improvement. This is very important for the following researchers who pay attention to our researches.

Several detection results of YOLOv5-l are presented in Fig. 8. The white boxes are the missed detected anomalies. They have some characteristics, such as small scale, heterogeneous background interference. In view of these difficulties, some solutions are provided for reference. For small-scale anomaly, improving the input resolution of the network [31] or multi-scale feature fusion strategy will promote the detection accuracy. For heterogeneous background interference, filter-based image preprocessing approach [5] or attention mechanism [6] can be used to suppress the background disturbance and

highlight the defect region.

V. CONCLUSION

This paper builds the photovoltaic solar cell electroluminescence image anomaly detection dataset (PVEL-AD), a novel dataset for anomaly detection in open-world industrial anomaly inspection scenarios. The proposed dataset is collected from real-world industrial solar-cell production lines, rather than lab environment. In the future, more samples with expert annotations will be collected into this dataset. It can be employed to evaluate several vision-based tasks, such as anomaly detection, few-shot anomaly detection, one-class classification and anomaly generation. Several state-of-the-art methods for object detection are thoroughly evaluated on this dataset. The experimental evaluations provide the first benchmark on this dataset, and show that there is still considerable room for improvement. Here are some improvement methods for reference, such as using rich defect-free samples to augment the defective dataset, employing a more expressive feature extraction network, introducing an attention mechanism, or applying cascade strategy for prediction, etc. Moreover, clean energy photovoltaic cells are very popular

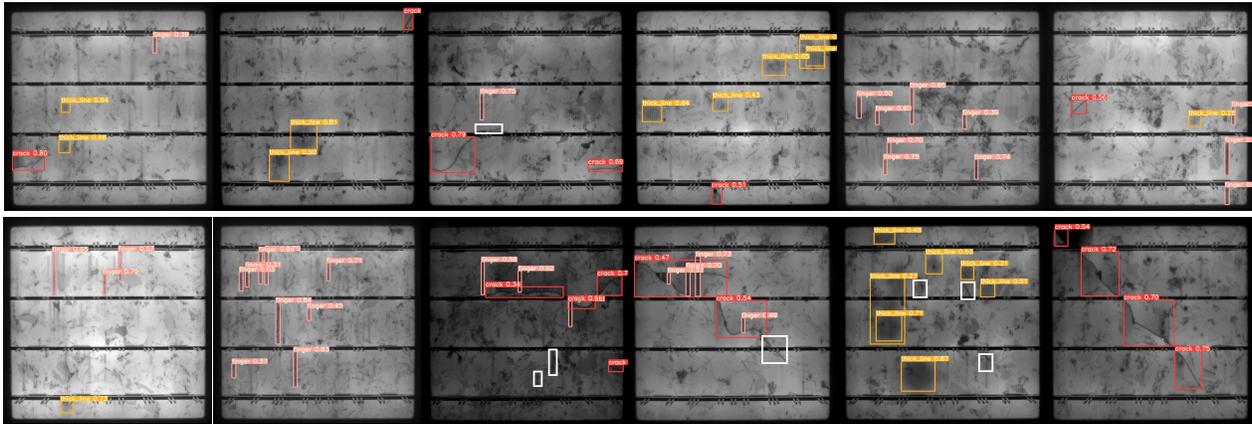


Fig. 8: Anomaly detection results of YOLOv5l.

recently. A large and comprehensive public dataset such as our PVEL-AD dataset will attract more scholars to participate in photovoltaic fault diagnosis, and then promoting the production quality of photovoltaic cells.

The limitation is that EL imaging technology requires electrical contacts, which poses an additional risk to the solar cell. An alternative option is non-contact photoluminescence (PL) imaging technology [35], which has no damage to the solar cell and can measure multiple times. Moreover, when the training sample is small, F-labelImg will degenerate into labelImg, because over-fitting may occur, which leads to failure of the anomaly detection.

APPENDIX

We conduct the comprehensive evaluation of all detection categories (8 types) in our PVEL-AD dataset, which are presented in Table VIII. The experimental setting is same as Section IV. A. The distribution of training data is illustrated in Table IV, which is extremely unbalanced. However, these algorithms such as EfficientDet and YOLOv5 still can obtain relatively stable experimental results (mAP50), which again verifies the good annotation quality of our PVEL-AD dataset.

ACKNOWLEDGMENT

First, the authors would like to thank the annotators who accelerate the release of PVEL-AD dataset: Rong Lei (M.S.) and Jiaqi Zhang (M.S.). Second, the authors would like to thank those workers who have pay attention to collect the dataset in Yingli Co., Ltd.

REFERENCES

[1] D. Tsai, S. Wu and W. Chiu, "Defect Detection in Solar Modules Using ICA Basis Images," *IEEE Trans. Ind. Inform.*, vol. 9, no. 1, pp. 122–131, Feb. 2013.

[2] B. Su, H. Chen, Y. Zhu, W. Liu and K. Liu, "Classification of Manufacturing Defects in Multicrystalline Solar Cells With Novel Feature Descriptor," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 12, pp. 4675–4688, Dec. 2019.

[3] K. Liu, H. Yan, and K. Meng, "Iterating Tensor Voting: A Perceptual Grouping Approach for Crack Detection on EL Images," *IEEE Trans. Autom. Sci. Eng.*, vol. 18, no. 2, pp. 831-839, Apr. 2021.

[4] B. Su, H. Chen, and P. Chen, "Deep Learning-Based Solar-Cell Manufacturing Defect Detection With Complementary Attention Network," *IEEE Trans. Ind. Inform.*, vol. 17, no. 6, pp. 4084–4095, Jun. 2021.

[5] H. Chen, H. Zhao, D. Han, and W. Liu, "Accurate and robust crack detection using steerable evidence filtering in electroluminescence images of solar cells," *Opt. Lasers Eng.*, vol. 118, no. 19, pp. 22-33, Jul. 2019.

[6] 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.

[7] T. Fuyuki and A. Kitiyanan, "Photographic diagnosis of crystalline silicon solar cells utilizing electroluminescence," *Appl. Phys. A*, vol. 96, no.7, pp. 189–196, Jul. 2009.

[8] K. Song, and Y. Yan, "A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects," *Appl. Surf. Sci.*, vol. 285, no. 21, pp. 858-864, Nov. 2013.

[9] L. Li, W. Ma, L. Li, and L. Cheng, "Research on Detection Algorithm for Bridge Cracks Based on Deep Learning," *Acta Automatica Sinica*, vol. 45, no. 9, pp. 1727–1742, Sep. 2019.

[10] D. Carrera, F. Manganini, G. Boracchi and E. Lanzarone, "Defect Detection in SEM Images of Nanofibrous Materials," *IEEE Trans. Ind. Inform.*, vol. 13, no. 2, pp. 551–561, Apr. 2017.

[11] D. Tabernik, S. Sela, J. Skvarc and D. Skocaj, "Segmentation-Based Deep-Learning Approach for Surface-Defect Detection," *J. Intell. Manuf.*, vol. 31, no. 5, pp. 759–776, May. 2019.

[12] P. Bergmann, M. Fauser, D. Sattlegger and C. Steger, "MVTec AD: A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 9584–9592.

[13] C. Buerhop, S. Deitsch, and A. Maier, "A benchmark for visual identification of defective solar cells in electroluminescence imagery," in *Proc. Eur. PV Sol. Energy Conf. Exhibit.*, 2018, pp. 1287–1289.

[14] S. Deitsch, V. Christlein, and S. Berger. "Automatic classification of defective photovoltaic module cells in electroluminescence images," *Sol. Energy*, vol. 185, no. 5, pp. 455–468, May. 2019.

[15] M. Akram, G. Li, and Y. Jin, "CNN based automatic detection of photovoltaic cell defects in electroluminescence images," *Energy*, vol. 189, no. 10, pp. 1–15, Oct. 2019.

[16] X. Qian, J. Li, J. Cao, Y. Wu, and W. Wang, "Micro-cracks detection of solar cells surface via combining short-term and long-term deep features," *Neural Netw.*, vol. 127, no. 7, pp. 132-140, Jul. 2020.

[17] C. Ge, Z. Liu, L. Fang, H. Ling, A. Zhang and C. Yin, "A Hybrid Fuzzy Convolutional Neural Network Based Mechanism for Photovoltaic Cell Defect Detection With Electroluminescence Images," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 7, pp. 1653-1664, Jul. 2021.

[18] H. Chen, Y. Pang, and Q. Hu. "Solar cell surface defect inspection based on multispectral convolutional neural network," *J. Intell. Manuf.*, vol. 31, no. 2, pp. 453–468, Feb. 2020.

[19] H. Chen, Q. Hu, B. Zhai, H. Chen, and K. Liu, "A robust weakly supervised learning of deep Conv-Nets for surface defect inspection," *Neural Comput. Appl.*, vol. 32, no. 3, pp. 11229–11244, Mar. 2020.

[20] B. Du, Y. He, Y. He, J. Duan and Y. Zhang, "Intelligent Classification of Silicon Photovoltaic Cell Defects Based on Eddy Current Thermography and Convolution Neural Network," *IEEE Trans. Ind. Inform.*, vol. 16, no. 10, pp. 6242-6251, Oct. 2020.

TABLE VIII: Experimental Detection Results of 8 Categories in PVEL-AD Dataset.

Methods	AP50								mAP50
	line_crack	finger	black_core	thick_line	star_crack	corner	fragment	scratch	
Faster RPN-CNN [4]	59.69	88.37	90.77	68.67	68.04	28.49	52.07	0	57.01
BAF-Detector [6]	62.05	88.47	90.85	68.71	73.11	27.69	52.53	0	57.93
EfficientDet-D0 [31]	55.30	17.77	98.41	70.89	69.60	0.04	0.01	0	39.00
EfficientDet-D1 [31]	55.76	50.24	98.55	73.99	71.94	2.81	0.46	0	44.22
EfficientDet-D2 [31]	54.46	53.37	76.26	76.06	73.14	0.06	0.02	8.33	42.21
YOLOv5-s [32]	60.79	92.03	98.07	65.53	59.23	31.67	85.26	33.33	65.74
YOLOv5-m [32]	57.95	90.51	97.52	64.72	63.18	17.75	68.01	33.33	61.62
YOLOv5-l [32]	59.27	91.31	97.76	66.04	62.84	16.81	54.71	33.33	60.26

[21] H. Han, C. Gao, and Y. Zhao, "Polycrystalline silicon wafer defect segmentation based on deep convolutional neural networks," *Pattern Recogn. Lett.*, vol. 130, no. 2, pp. 234-241, Feb. 2020.

[22] M. R. U. Rahman and H. Chen, "Defects Inspection in Polycrystalline Solar Cells Electroluminescence Images Using Deep Learning," *IEEE Access*, vol. 8, no.2, pp. 40547-40558, Feb. 2020.

[23] B. Kang, Z. Liu, X. Wang, F. Yu, J. Feng and T. Darrell, "Few-Shot Object Detection via Feature Reweighting," in *Proc. Int. Conf. Comput. Vis.*, 2019, pp. 8419-8428.

[24] L. Qiao, Y. Zhao, and Z. Li, "DeFRCN: Decoupled Faster R-CNN for Few-Shot Object Detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2021, pp. 8661-8670.

[25] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: A retrospective," *Int. J. Comput. Vis.*, vol. 111, no. 1, pp. 98-136, Jan. 2015.

[26] P. Bergmann, M. Fauser, D. Sattlegger and C. Steger, "Uninformed Students: Student-Teacher Anomaly Detection With Discriminative Latent Embeddings," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 4182-4191.

[27] S. Niu, B. Li, X. Wang and H. Lin, "Defect Image Sample Generation With GAN for Improving Defect Recognition," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 3, pp. 1611-1622, Jul. 2020.

[28] B. Su, Z. Zhou, H. Chen, and X. Cao, "SIGAN: A Novel Image Generation Method for Solar Cell Defect Segmentation and Augmentation," May. 2021. [Online]. Available: <https://arxiv.org/abs/2104.04953>

[29] S. Yun, D. Han, S. Chun, S. J. Oh, Y. Yoo and J. Choe, "CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2019, pp. 6022-6031.

[30] C. Wang, H. Mark Liao, Y. Wu, P. Chen, J. Hsieh and I. Yeh, "CSPNet: A New Backbone that can Enhance Learning Capability of CNN," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 1571-1580.

[31] M. Tan, R. Pang and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 10778-10787.

[32] G. Jocher, A. Stoken, and J. Borovec, "ultralytics/yolov5: v3.1 - Bug Fixes and Performance Improvements," Oct. 2020. [Online]. Available: <https://doi.org/10.5281/zenodo.4154370>

[33] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Learn. Represent.*, 2019, pp. 10691-10700.

[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770-778.

[35] M. Demant, M. Glatthaar, J. Haunschild, and S. Rein, "Analysis of luminescence images applying pattern recognition techniques," in *Proc. Eur. PV Sol. Energy Conf. Exhibit.*, pp. 1078-1082, Sep. 2010.



Binyi Su received the B.S. degree in intelligent science and technology from the Hebei University of Technology, Tianjin, China, in 2017, and the M.S degree in control engineering from the Hebei University of Technology, Tianjin, China, in 2020.

He is currently pursuing the Ph.D. degree in the State Key Laboratory of Virtual Reality Technology and Systems, computer science and technology from Beihang University, Beijing, China. His current research interests include computer vision and pattern recognition, anomaly detection, and smart city.



Zhong Zhou received the B.S. degree in material physics from Nanjing University in 1999 and the Ph.D. degree in computer science and technology from Beihang University, Beijing, China, in 2005.

He is a Professor and the Ph.D. Adviser with the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University. His main research interests include virtual reality, augmented reality, computer vision, and artificial intelligence.



Haiyong Chen received the M.S. degree in detection technology and automation from the Harbin University of Science and Technology, Harbin, China, in 2005, and the Ph.D. degree in control science and engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2008.

He is a Professor with the School of Artificial Intelligence and Data Science, Hebei University of Technology, Tianjin. His current research interests include image processing, robot vision.