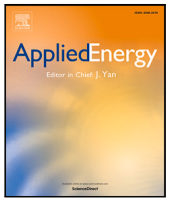
Applied Energy 372 (2024) 123759

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy



Fast object detection of anomaly photovoltaic (PV) cells using deep neural networks

Jinlai Zhang [a], Wenjie Yang [b,*], Yumei Chen [c], Mingkang Ding [e], Huiling Huang [d], Bingkun Wang [f], Kai Gao [a], Shuhan Chen [a], Ronghua Du [a]

[a] College of Automotive and Mechanical Engineering, Changsha University of Science and Technology, Changsha, 410114, Hunan, China
[b] College of Mathematics and Statistics, Hu'nan University of Finance and Economics, Changsha, 410205, Hunan, China
[c] School of Mechanical Engineering, Hunan University of Science and Technology, Xiangtan, 411201, Hunan, China
[d] School of Mechanical Engineering, Xiangtan University, Xiangtan, 411105, Hunan, China
[e] School of Information Science and Technology, Taishan University, Taian, 271021, Shandong, China
[f] International College of Engineering, Changsha University of Science and Technology, Changsha, 410114, Hunan, China

ARTICLE INFO

Keywords:
DNNs
Solar energy
Computer vision

ABSTRACT

Anomaly detection in photovoltaic (PV) cells is crucial for ensuring the efficient operation of solar power systems and preventing potential energy losses. In this paper, we propose an enhanced YOLOv7-based deep learning framework for fast and accurate anomaly detection in PV cells. Our approach incorporates Partial Convolution, Switchable Atrous Convolution and novel data augmentation techniques to address the challenges of varying defect sizes, complex backgrounds. The Partial Convolution component effectively manages the irregularities in PV cell images, reducing false detections. On the other hand, the Switchable Atrous Convolution enhances the model's adaptability to different defect scales and spatial resolutions, leading to improved localization and classification performance. We evaluate our model on a large-scale dataset of PV cell images, demonstrating its superiority over existing methods in terms of detection accuracy, speed, and robustness. Our proposed framework offers a practical and reliable solution for real-time anomaly detection in PV cells, facilitating timely maintenance and maximizing the performance of solar energy systems. The integration of these advanced convolution techniques into the YOLOv7 model not only improves detection capabilities but also paves the way for further research and development in the field of deep learning-based anomaly detection.

1. Introduction

Solar energy [1,2] has emerged as one of the most promising alternatives to traditional fossil fuels, owing to its abundance, sustainability, and clean nature [3,4]. Photovoltaic (PV) cells, which convert sunlight into electricity, play a pivotal role in harnessing solar energy [5]. As the demand for solar power systems grows globally, ensuring the optimal performance and longevity of PV cells becomes increasingly important. Anomaly detection, the process of identifying defects or malfunctions in PV cells, is a critical aspect of this endeavor, as it allows for timely maintenance and prevents potential energy losses.

Recent years have witnessed significant advancements in deep learning techniques [6,7], particularly in the domain of computer vision [8–10]. These developments have led to the application of deep learning-based methods for anomaly detection in PV cells. Convolutional Neural Networks (CNNs) [11,12] have demonstrated exceptional capabilities in identifying patterns and features in image data, making them suitable for detecting defects in PV cell images. Among the various deep learning architectures, the You Only Look Once (YOLO) family of models [13,14] has gained considerable attention for their real-time object detection performance.

Anomaly detection in PV cells has been studied extensively in recent years [9,15]. Traditional methods, such as statistical analysis, signal processing, and machine learning algorithms, have been employed for defect detection. However, these approaches often suffer from limitations in terms of accuracy, adaptability, and computational efficiency. With the advent of deep learning techniques, researchers have explored the use of CNNs for PV cell anomaly detection [15]. While these studies have shown promise, they often struggle to handle the complexities of real-world PV cell images, such as varying defect sizes, irregular shapes, and diverse backgrounds. The YOLO family of models has been widely adopted for object detection tasks due to its impressive speed and

* Corresponding author.
E-mail address: gs_ywj@163.com (W. Yang).

https://doi.org/10.1016/j.apenergy.2024.123759
Received 31 August 2023; Received in revised form 29 November 2023; Accepted 17 June 2024
Available online 1 July 2024

accuracy. YOLOv7 [14], the latest iteration in the series, has demonstrated state-of-the-art (SOTA) performance across various benchmark datasets. However, its application to PV cell anomaly detection remains limited, and there is a need to further optimize the model for this specific context.

Our research is motivated by the need for a fast and accurate deep learning-based approach to anomaly detection in PV cells that can address the challenges posed by real-world image data. The integration of Partial Convolution and Switchable Atrous Convolution techniques into the YOLOv7 model aims to improve its adaptability to varying defect scales, spatial resolutions, and data irregularities. By doing so, we seek to contribute to the ongoing efforts in the field of deep learning-based anomaly detection and facilitate the development of more efficient solar energy systems.

The remainder of this paper is organized as follows: Section 2 presents the literature review. Section 3 presents the methodology, detailing the YOLOv7 architecture and the proposed PSA-YOLOv7. Section 4 describes the dataset used for model evaluation and the performance metrics employed for comparison. Section 4 reports the experimental results, demonstrating the effectiveness of our proposed framework in comparison to existing methods. Section 5 discusses the implications of our findings and the potential avenues for future research. Section 6 concludes the paper, summarizing our contributions and their significance in the context of PV cell anomaly detection.

2. Literature review

2.1. Traditional methods for PV cell anomaly detection

Before the emergence of deep learning techniques, various traditional methods were employed for anomaly detection in photovoltaic (PV) cells. These methods can be broadly categorized into two groups: statistical analysis, and signal processing. In this section, we provide a comprehensive review of these traditional approaches, highlighting their strengths and limitations in detecting defects in PV cells.

2.1.1. Statistical analysis

Statistical analysis methods rely on the mathematical properties of the data to identify anomalies. Common techniques used for PV cell anomaly detection include hypothesis testing, regression analysis, and control charts. Hypothesis testing is used to assess whether the observed data deviates significantly from the expected behavior, suggesting the presence of anomalies. Fouzi et al. [16] developed a hypothesis testing method to detect anomalies in PV plants. Vergura Silvano [17] proposes a methodology to monitor the energy performance of photovoltaic plants statistically, without using environmental parameters. [18] proposes a machine learning approach using Gaussian process regression (GPR) and a generalized likelihood ratio test (GLRT) chart to enhance fault detection performance in photovoltaic (PV) systems.

While statistical analysis methods are relatively simple and computationally efficient, they often suffer from several limitations. They rely on strong assumptions about the data distribution, which may not always hold in real-world scenarios. Although deep learning also suffers from the same problem, data augmentation techniques can reduce the impact of this problem. Additionally, they may fail to capture complex relationships between variables and struggle to handle high-dimensional data, leading to suboptimal detection performance.

2.1.2. Signal processing

Signal processing techniques have also been applied to PV cell anomaly detection. These methods involve transforming the data into an alternative representation to extract relevant features for defect identification [19]. Some commonly used signal processing techniques include Fourier Transform and Principal Component Analysis (PCA). Fourier Transform decomposes a signal into its constituent frequencies, enabling the identification of abnormal frequency components associated with defects. Harrou, et al. [20] presents a model-based anomaly detection method using k-nearest neighbors (kNN) and univariate monitoring approaches with parametric and nonparametric thresholds for direct current (dc) side fault detection in photovoltaic (PV) systems. PCA, on the other hand, reduces the dimensionality of the data by projecting it onto a lower-dimensional subspace that captures most of its variance. By doing so, PCA can reveal hidden patterns and structures that may be indicative of anomalies. Arena et al. [21] proposes a robust anomaly detection method for the photovoltaic production factory scenario using Monte Carlo based pre-processing, principal component analysis, and key performance indicators to isolate anomalous conditions and trigger an alarm when exceeding a reference threshold.

Despite their ability to extract useful features from the data, signal processing methods have some drawbacks. They often require expert knowledge to select appropriate parameters and transform techniques, which can be time-consuming and prone to errors. Furthermore, these methods may be sensitive to noise and susceptible to false detections.

In summary, traditional methods for PV cell anomaly detection, including statistical analysis, signal processing, and machine learning algorithms, have made substantial contributions to the field. However, they often face challenges when dealing with the complexities of real-world PV cell data, such as varying defect sizes, irregular shapes, diverse backgrounds, and high-dimensional data. Moreover, these methods usually require significant domain expertise and extensive parameter tuning, which can be time-consuming and error-prone. At the same time, statistical analysis can be complementary to deep learning methods.

The advent of deep learning techniques, particularly Convolutional Neural Networks (CNNs), has opened up new possibilities for anomaly detection in PV cells. Although deep learning also requires experts to label data, these models have demonstrated exceptional capabilities in learning hierarchical representations of image data and have shown great potential in addressing the limitations of traditional methods. In the following sections, we review the literature on CNN-based approaches for PV cell anomaly detection and introduce the YOLO family of models, which form the basis of our proposed framework.

2.2. CNN-based approaches for anomaly detection

With the rapid development of deep learning techniques, Convolutional Neural Networks (CNNs) [11,12] have emerged as a powerful tool for various computer vision tasks, including image classification, object detection, and segmentation [22]. Their ability to automatically learn hierarchical features from image data has led to their application in anomaly detection for photovoltaic (PV) cells. In this section, we provide a comprehensive review of the literature on CNN-based approaches for PV cell anomaly detection, discussing their advantages and limitations.

The fundamental building blocks of a CNN are convolutional layers, pooling layers, and fully connected layers [11]. Convolutional layers perform spatial feature extraction using a set of learnable filters, which are convolved with the input image to produce feature maps. Pooling layers reduce the spatial dimensions of the feature maps, allowing for translation invariance and reducing the number of parameters. Fully connected layers are used to map the extracted features to the output, such as class labels or bounding boxes for object detection.

Various CNN architectures have been proposed in the literature for PV cell anomaly detection [23,24]. These architectures can be broadly classified into two categories: classification and detection networks. Classification networks, such as AlexNet, VGG, and ResNet, consist of a series of convolutional and pooling layers followed by fully connected layers. These models have been employed for binary or multi-class classification of PV cell defects, with the input being a cropped image or a fixed-size patch centered around the defect. Early

work mostly classified photovoltaic panels into two categories: normal and abnormal [24,25].

Detection Networks have gradually gained recognition in academia because they can locate and box in the specific location of abnormalities [8,9]. More advanced CNN-based models are detection networks, such as Faster R-CNN, and YOLO, have been employed for anomaly detection in PV cells by leveraging their object detection capabilities [9]. Among them, the You Only Look Once (YOLO) family of models has gained significant attention in recent years due to their real-time object detection capabilities and exceptional performance across a wide range of applications. Given their exceptional performance in various object detection tasks, YOLO models have been explored for anomaly detection in PV cells as well [9]. Their ability to perform real-time object detection and localization makes them suitable for large-scale PV cell monitoring systems, where timely identification of defects is critical for maintaining optimal efficiency.

Researchers have adapted YOLO models for PV cell anomaly detection by training them on annotated datasets containing various types of defects, such as cracks, hotspots, and soiling. The models have demonstrated promising results in terms of detection accuracy, speed, and robustness, outperforming traditional methods and other CNN-based approaches in many cases.

Despite their impressive performance, YOLO models can still be further improved to better handle the complexities and challenges associated with PV cell anomaly detection. In this paper, we propose a novel framework that enhances the YOLOv7 model by incorporating Partial Convolution and Switchable Atrous Convolution techniques, which address issues related to irregularities in PV cell images and varying defect scales.

3. Methodology

In this paper, we propose a novel approach for fast anomaly detection of PV cells. In this section, we first present the overview of YOLOv7 [14] and the proposed PSA-YOLOv7, and then we introduce each component within PSA-YOLOv7.

3.1. Overview of the YOLOv7 architecture

Inspired by some recent work in computer vision [9], we chose YOLOv7 [14] as the baseline model. YOLOv7, the latest iteration in the YOLO family of models, builds upon the success of its predecessors by incorporating advanced techniques and architectural improvements to achieve state-of-the-art object detection performance. In this section, we provide a comprehensive overview of the YOLOv7 architecture, discussing its key components, the underlying design principles, and the innovations that differentiate it from previous versions.

YOLOv7 is a state-of-the-art object detection model that is trained to detect objects from an image or video with high accuracy and speed. It is an extension of the You Only Look Once (YOLO) family [26] of object detection models. As shown in Fig. 1, the architecture of YOLOv7 consists of a backbone network and a head network, the details of the CBS, MPC and ELAN module can refer to [14]. The backbone network extracts features from the input image, which are then passed through the neck network to further process the features and reduce their dimensionality. The ELAN module plays a critical role in the backbone. Finally, the head network uses these features to predict the bounding boxes and class probabilities of objects in the image. The output of YOLOv7 is a set of bounding boxes, each of which represents an object in the image. The bounding boxes are represented as a set of four coordinates: (x, y) for the top-left corner of the box, and (w, h) for the width and height of the box, respectively. The class probabilities for each bounding box are also predicted by the model. These probabilities represent the likelihood of the object being of a certain class, and are represented as a vector of probabilities for each class.

The loss function used in YOLOv7 is a combination of several terms, including the bounding box regression loss, the objectness loss, and the class probability loss. The loss function is optimized using stochastic gradient descent (SGD) [27] with backpropagation.

The overall objective of YOLOv7 is to minimize the following loss function:

$$\mathcal{L} = \sum_{i=1}^{S^2} \sum_{j=1}^{B} \left[\lambda_{obj}^{ij} \left(L_{coord}^{ij} + L_{conf}^{ij} \right) + \sum_{c=1}^{C} \lambda_{class}^{ij} L_{class}^{ij,c} \right], \tag{1}$$

where S is the number of grid cells, B is the number of bounding boxes per cell, C is the number of classes, λ_{obj}, λ_{class} are the weights for the corresponding terms, L_{coord}, L_{conf}, and L_{class} are the loss terms for bounding box regression, objectness, and class probability, respectively.

3.2. Overview of the proposed PSA-YOLOv7 architecture

Although YOLOv7 has shown good performance in anomaly detection of PV cells, it still has the following defects:

Detection accuracy: While YOLOv7 demonstrates impressive object detection capabilities in general, its performance in detecting anomalies in PV cells, particularly those with weak color cracks or subtle defects, may be limited. Improving YOLOv7 for PV cell anomaly detection ensures better identification of various defects, leading to more reliable monitoring and maintenance of solar panels.

Adaptability to defect scale: PV cell defects can span a wide range of scales, from small micro-cracks to large-scale delamination. The existing YOLOv7 architecture may not be optimal for handling defects of varying sizes. Enhancing YOLOv7 with adaptive techniques can improve the model's ability to detect defects across different scales.

Robustness to varying conditions: PV cell installations are subjected to diverse environmental conditions, such as varying lighting, weather, and occlusions. The YOLOv7 architecture may not be robust enough to handle these variations effectively, leading to suboptimal detection performance. Improving YOLOv7 can increase its robustness, ensuring reliable anomaly detection across various conditions.

Handling irregular defects and missing data: PV cell images may contain irregular defects or missing data due to occlusions, sensor noise, or other factors. The standard YOLOv7 architecture may struggle to handle these complexities, resulting in reduced detection accuracy. Incorporating advanced techniques, such as partial convolution or switchable atrous convolution, can enhance YOLOv7's ability to process irregular defects and missing data effectively.

Real-time detection capabilities: Fast anomaly detection is crucial for large-scale PV cell monitoring systems, where timely identification and remediation of defects are essential for maximizing energy production and minimizing downtime. Improving YOLOv7 to maintain real-time detection capabilities while enhancing its performance in PV cell anomaly detection ensures that the framework remains suitable for real-world solar energy installations.

In order to improve the performance of YOLOv7 in the above situations, in this paper, we propose **PSA-YOLOV7**. The structure of PSA-YOLOV7 is shown in Fig. 2, and it has made improvements as shown in Table 1 relative to YOLOV7. Specifically, in order to make the improved model achieve real-time detection performance and reduce the number of parameters of the model, we introduced partial convolution [28] in the ELAN module of the backbone part. In order to improve YOLOV7 under different defect scales, irregular defects and missing data, we introduced switchable atrous convolution [29] to the backbone part. To improve the robustness to varying conditions, we propose novel data augmentation methods. In the following, we introduce each component in detail.

3.3. Partial ELAN module

Traditional convolution operations in CNNs assume that the input data is uniformly sampled and complete [30], which may not be the

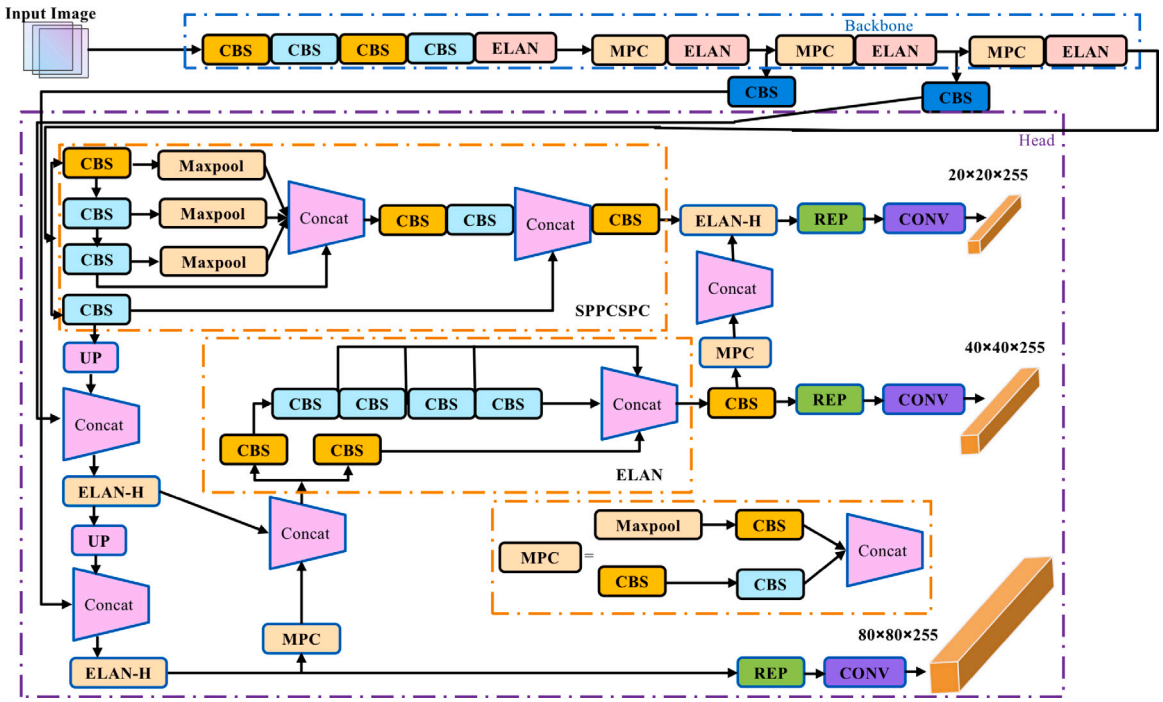


Fig. 1. Detailed architecture of YOLOV7.

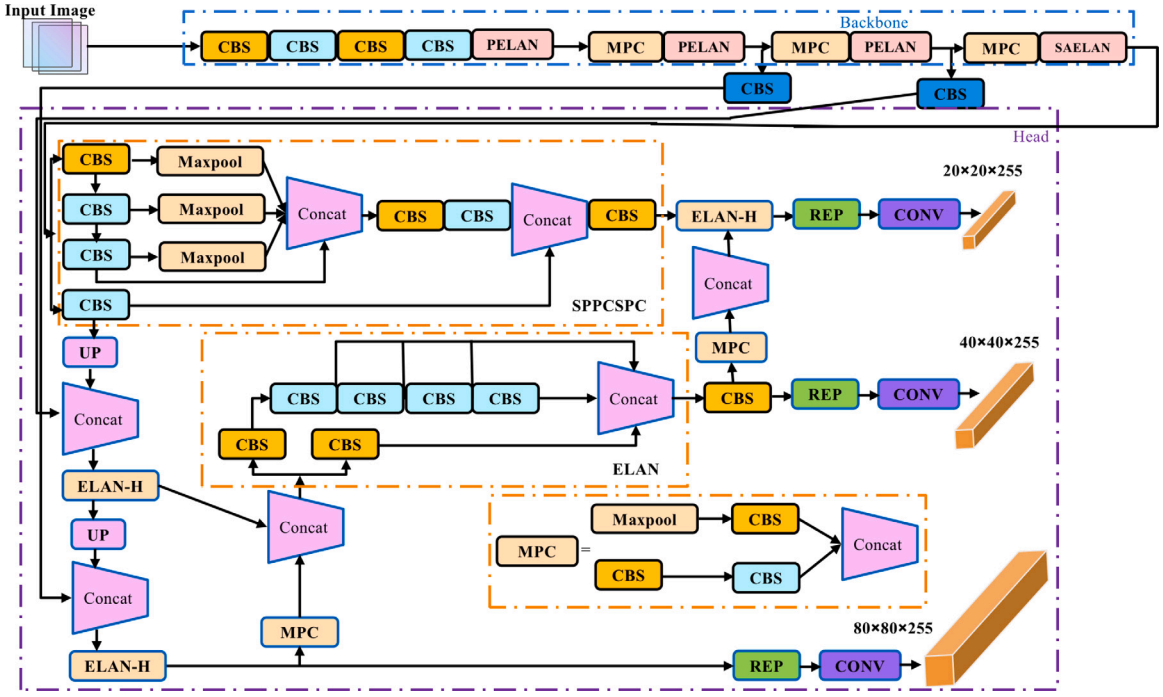


Fig. 2. Detailed architecture of PSA-YOLOV7.

Table 1
Structural comparisons between YOLOv7 and PSA-YOLOv7.

Layers	YOLOv7	PSA-YOLOv7
C5	ELAN	PELAN
C7	ELAN	PELAN
C9	ELAN	PELAN
C11	ELAN	SAELAN

case for PV cell images with irregular defects or occlusions. When dealing with such images, standard convolutions can result in sub-optimal feature extraction and reduced detection accuracy. Moreover, doing complete convolution operations may also introduce redundant parameters and computations. In this context, to enable the enhanced model to attain real-time detection performance while simultaneously reducing the model's parameter count, we incorporate partial convolution [28] into the ELAN module of the backbone component. In order to minimize the model parameters, we have improved the ELAN module

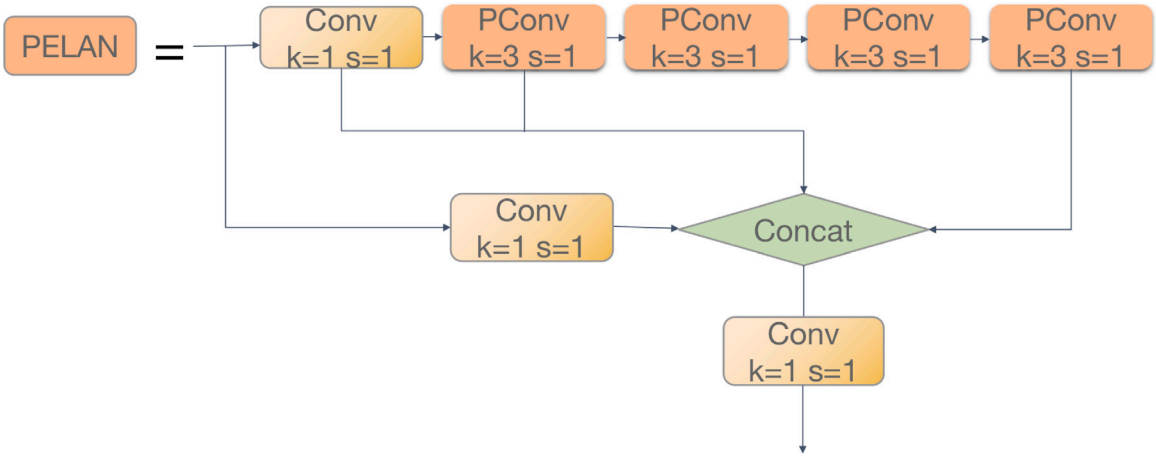


Fig. 3. Detailed architecture of PELAN.

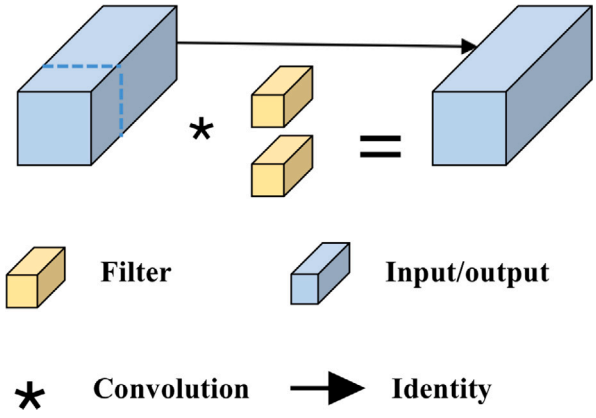


Fig. 4. Detailed architecture of Partial Convolution.

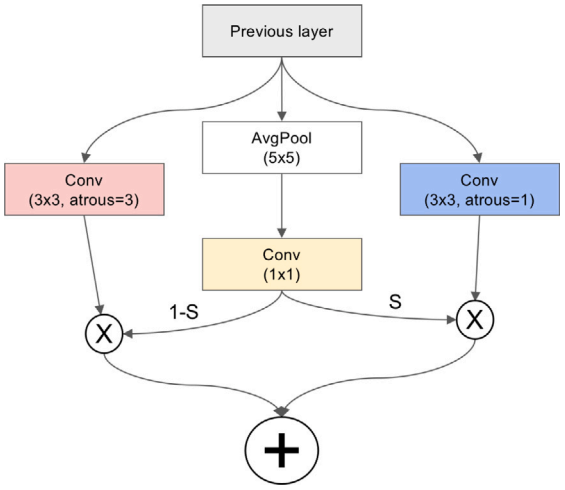


Fig. 5. Detailed architecture of Switchable Atrous Convolution.

in the backbone part, and finally formed the Partial ELAN module proposed in this paper, referred to as PELAN module, as shown in Fig. 3.

The partial convolution (PConv) [28] reduces computational redundancy by applying filters on only a few input channels while leaving the remaining ones untouched. PConv obtains lower floating-point operations per second (FLOPs) than the regular convolution and higher FLOPS than the depthwise/group convolution [31]. The formula for partial convolution can be represented as follows:

$$y_{i,j} = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} w_{k,l} x_{i+k,j+l} \cdot m_{i+k,j+l}$$
$$m_{i,j} = \begin{cases} 1 & \text{if } x_{i,j} \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where x is the input tensor, w is the convolution kernel, m is a binary mask that indicates which pixels are valid in x, and y is the output tensor. The detailed architecture of partial convolution is shown in Fig. 4, where the part enclosed by the blue dotted line represents the input channel selected by the binary mask.

To integrate PConv into the YOLOv7 architecture, we replace the standard convolution layers in the backbone. This modification allows the model to better handle irregular defects and missing data in the PV cell images, resulting in improved feature extraction and anomaly detection performance. The integration of Partial Convolution into YOLOv7 does not significantly affect the computational complexity of the model, allowing it to maintain real-time detection capabilities. We changed the convolutional layer of the ELAN module, and replaced all four of them with PConv, and the resulting PELAN module is shown in Fig. 3.

The integration of Partial Convolution into YOLOv7 is seamless and can reduce the model's computational complexity, allowing it to maintain its real-time detection capabilities.

3.4. Switchable atrous ELAN module

Traditional convolution operations in CNNs [32] have a fixed receptive field size [33], which may not be suitable for detecting objects or defects of various scales in PV cell images. Although multi-scale feature extraction techniques, such as those employed in the YOLOv7 architecture, alleviate this issue to some extent, they may still suffer from suboptimal performance when dealing with a wide range of defect scales. Switchable Atrous Convolution [29] aims to address this issue by providing an adaptive and learnable convolution operation that adjusts its receptive field size based on the input data, resulting in more effective detection of defects across different scales. In this case, to enhance the performance of YOLOv7 when dealing with varying defect scales, irregular defects, and missing data of PV cells, we introduce switchable atrous convolution [29] to the backbone component.

Switchable Atrous Convolution (SAConv) [29] convolves the same input feature with different atrous rates and gathers the results using switch functions. SAConv improves the detector performance by a large margin by converting all the standard 3 × 3 convolutional layers

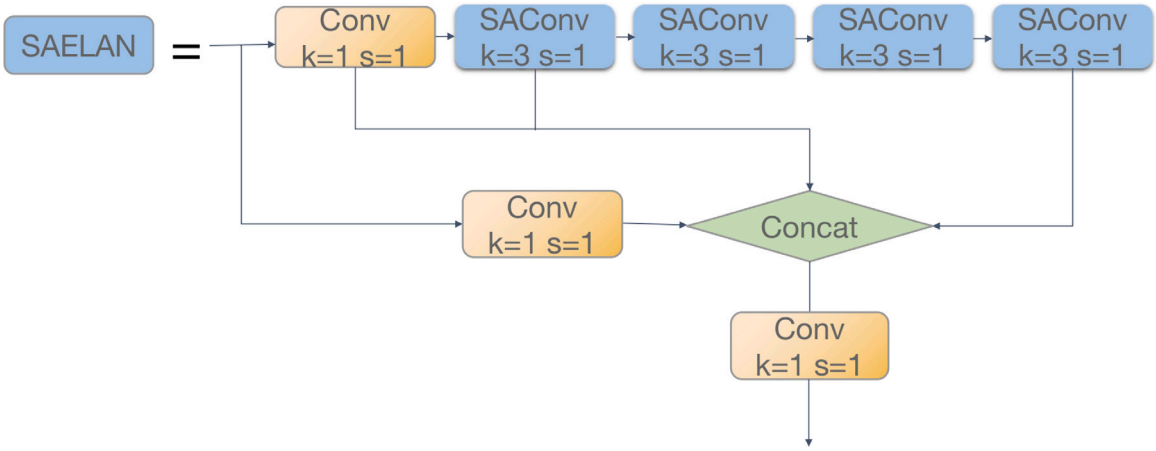


Fig. 6. Detailed architecture of SAELAN.

in the bottom-up backbone to SAConv. SAConv provides a mechanism to easily convert pre-trained standard convolutional networks (e.g., ImageNet-pretrained checkpoints). A new weight locking mechanism is used in SAConv where the weights of different atrous convolutions are the same except for a trainable difference. The concept of SAConv is illustrated in Fig. 5. The mathematical formula for SAConv can be expressed as follows:

$$y_i = \sum_{j=1}^{n} w_j \cdot x_{i+jr_j} \cdot s_{i+jr_j} \tag{3}$$

where y_i is the output at position i, x is the input feature map, w_j is the weight of filter j, r_j is the atrous rate of filter j, and s_{i+jr_j} is a binary switch function that controls whether to use filter j at position $i + jr_j$. Switchable Atrous Convolution is based on the concept of atrous (dilated) convolutions [34], which have a larger receptive field than standard convolutions by introducing gaps between the filter elements. In SAConv, multiple atrous convolution layers with different dilation rates are applied in parallel, and their outputs are adaptively combined using a learnable gating mechanism. This design allows the model to adjust its receptive field size based on the input data, enabling it to effectively capture defects and objects of varying sizes. The key advantage of Switchable Atrous Convolution is that it provides a more flexible and adaptive convolution operation that can handle a wide range of defect scales, resulting in improved detection performance in PV cell images.

To deal with a wide range of defect scales, we replace selected standard convolution layers in the backbone of YOLOv7 with SAConv layers. This modification allows the model to adapt its receptive field size based on the input data, resulting in improved detection performance for defects of varying sizes in the PV cell images. Similar to PELAN, we changed the convolutional layer of the ELAN module, and replaced four of them with SAConv, and the final Switchable Atrous ELAN (SAELAN) module is shown in Fig. 6.

3.5. RMosaic augmentation

Numerous empirical studies on data augmentation [35] have demonstrated their ability to bolster the robustness of models. Moreover, the production process of photovoltaic (PV) cells is frequently influenced by varying lighting conditions and environmental Gaussian noise. To strengthen the model's robustness in such situations, we propose a novel random contrast Mosaic data augmentation technique and dub our method RMosaic augmentation.

Mosaic data augmentation is a popular technique that combines multiple images into a single training example to improve model performance. In the series of work after YOLOv5, it is the default data augmentation configuration. This technique involves randomly selecting 4 images and blending them together to create a mosaic image. The mosaic image is randomly augmented using techniques such as scaling, flipping, and rotating. The resulting image is then used as a training example for the model.

The blending of the images is performed using the following formula:

$$\begin{bmatrix} I_1 & I_2 \\ I_3 & I_4 \end{bmatrix} \rightarrow [\text{Mosaic Image}] \tag{4}$$

where I_1, I_2, I_3, and I_4 are four randomly selected training images. The Mosaic Image is created by concatenating the four images together in a 2×2 grid and randomly adjusting their positions and sizes. This creates a larger, more diverse training image that helps the network learn to detect objects in different contexts and orientations. In addition to improving object detection performance, Mosaic data augmentation also helps reduce overfitting by increasing the effective size of the training dataset.

In our proposed random contrast Mosaic data augmentation (RMosaic), we add random contrast to each image. It can be presented as the following formula:

$$\begin{bmatrix} C(I_1) & C(I_2) \\ C(I_3) & C(I_4) \end{bmatrix} \rightarrow [\text{RMosaic Image}] \tag{5}$$

where $C()$ is the random contrast transformation, the RMosaic Image is augmented images during training.

4. Experimental results

4.1. Experimental setup

In this paper, we use Pytorch [36] to build the PSA-YOLOv7. To evaluate the performance of the proposed framework and compare it with existing methods, we conducted experiments on the PVEL-AD-2021 [9] benchmark, containing various types of defects such as linear crack, finger, black core and thick line. The dataset was split into training and test sets, ensuring a fair and unbiased comparison of the methods, and the number of training sets is 4604, and the number of test sets is 1622. The details of PVEL-AD-2021 used in this paper are shown in Table 2. All models are trained with 100 epochs on a Linux server with NVIDIA RTX 3090 GPU.

The compared methods include two-stage object detection models [8], and previous YOLO family models. All models were trained on the same training set and evaluated on the same test set using standard evaluation metrics, such as precision, recall, and mean Average Precision (mAP). Precision and recall are evaluation metrics used in classification tasks that measure how well a model performs in terms

Table 2
The details of PVEL-AD-2021 [9].

Class	train	test	total
linear crack	879	381	1260
finger	2175	782	2957
black core	818	210	1028
thick line	732	249	981

of correctly predicting positive and negative cases. In the context of anomaly detection, positive cases represent the anomalous photovoltaic (PV) cells, while negative cases are the normal PV cells. Precision is the proportion of true positives (TP) out of all predicted positive cases (TP + false positives (FP)):

$$Precision = TP/(TP + FP) \tag{6}$$

Recall, also known as sensitivity or true positive rate (TPR), is the proportion of true positives identified out of all actual positive cases:

$$Recall = TP/(TP + FN) \tag{7}$$

The mean Average Precision (mAP) is a widely used metric for the evaluation of object detection models [37]. It measures the accuracy and consistency of the model's predictions across various IoU (Intersection over Union) thresholds.

For each class, precision and recall are computed for all IoU thresholds. These values are then used to compute the Average Precision (AP), which is the area under the precision–recall curve. The mAP is the mean of the AP values over all classes. The formula for mAP is given as:

$$mAP = \frac{1}{n}\sum_{i=1}^{n} AP_i \tag{8}$$

where n is the total number of classes, and AP_i is the average precision of the i_{th} class.

By evaluating precision, recall, and mAP, we can assess the performance of our proposed PSA-YOLOv7 model in detecting anomalous PV cells accurately, consistently, and across multiple classes.

4.2. Comparison with the baseline model

In this section, we compare the proposed PSA-YOLOv7 with the baseline YOLOv7 model in terms of confusion matrix, precision, recall, and training loss curves. We demonstrate that our proposed model outperforms the baseline model in all aspects.

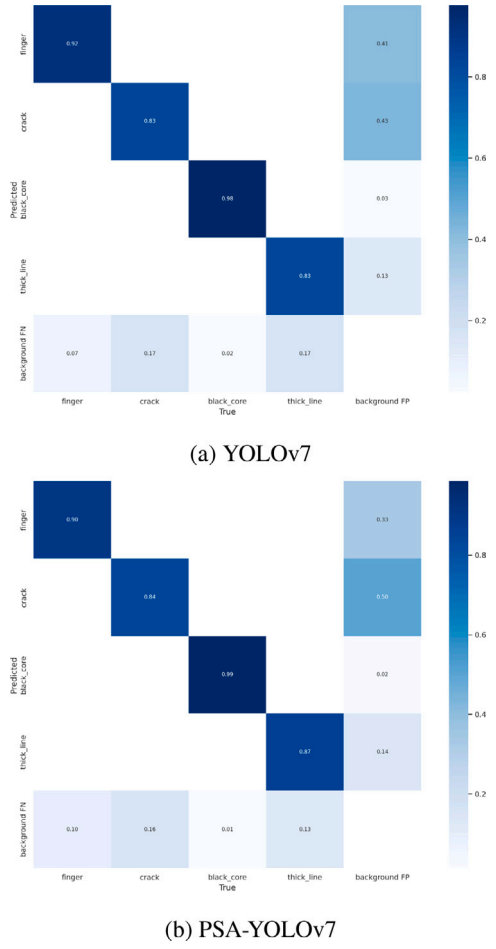
4.2.1. Confusion matrix

The confusion matrix is used to evaluate the performance of a classification model by visualizing the number of true positives, false positives, true negatives, and false negatives [38,39]. As shown in Fig. 7, the PSA-YOLOv7 model has a higher number of true positives than the baseline YOLOv7 model for crack, black core and thick line classes. This indicates that the proposed model is more accurate in identifying anomaly PV cells.

4.2.2. Precision and recall

Precision and recall are evaluation metrics used to measure the effectiveness of anomaly detection [24]. Precision measures the proportion of correct anomaly cells detected out of the total number of cells detected. Recall measures the proportion of correctly detected anomaly cells out of the total number of actual anomaly cells. Table 3 shows the precision and recall of the YOLOv7 and PSA-YOLOv7 models.

As shown in Table 3, the PSA-YOLOv7 model has a slightly lower recall than the baseline YOLOv7 model. However, the PSA-YOLOv7 model outperforms the baseline model in terms of precision. This indicates that the PSA-YOLOv7 model is more effective in identifying anomaly cells.



(a) YOLOv7

(b) PSA-YOLOv7

Fig. 7. The confusion matrix of (a) YOLOv7, (b) PSA-YOLOv7.

Table 3
Precision and recall of YOLOv7 and PSA-YOLOv7 models.

Model	Precision	Recall
YOLOv7	84.8	85.1
PSA-YOLOv7 (ours)	**88.3**	85.0

4.2.3. Training loss curves

Fig. 8 shows the training loss curves of the YOLOv7 and PSA-YOLOv7 models. The training loss curve shows the decrease in loss during the training process. The loss is calculated by the difference between the predicted values and actual values. As shown in Fig. 8, the PSA-YOLOv7 model has a lower box loss than the baseline YOLOv7 model, and has a higher and stable mAP than YOLOv7 model. This suggests that the proposed model is better optimized and converged.

4.3. Visualization

To visually demonstrate the differences between PSA-YOLOv7 and YOLOv7, we have visualized their anomaly detection results, as shown in Fig. 9. From the figure, we can see that YOLOv7 exhibits issues with missing detections and false positives, and its performance is inferior to the PSA-YOLOv7 proposed in this paper.

In addition, to further explore the advanced features of PSA-YOLOv7, we used the interpretability tool Grad-CAM [40] to visualize and analyze the anomaly detection results of PSA-YOLOv7 and YOLOv7. As shown in Fig. 10, the attention maps of PSA-YOLOv7 are mostly concentrated within the bounding box, while the attention maps of YOLOv7 are relatively scattered, with a significant portion of the

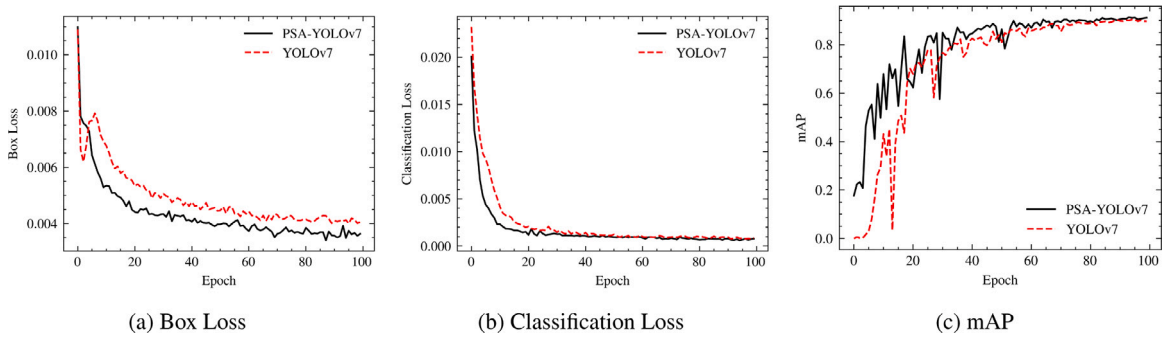


Fig. 8. The training curves of YOLOv7 and PSA-YOLOv7.

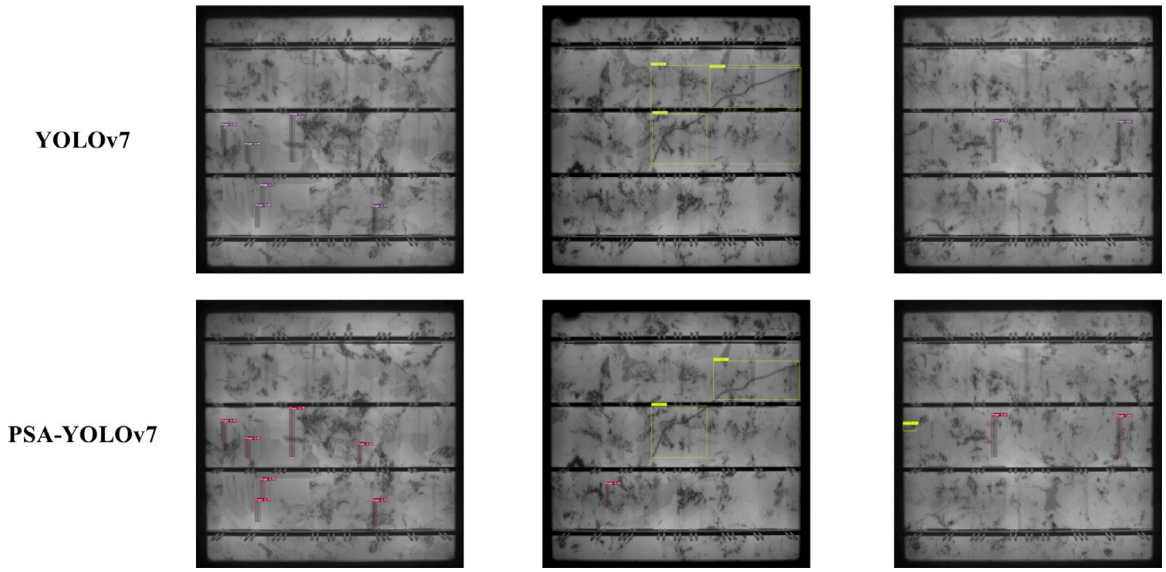


Fig. 9. The anomaly detection results of YOLOv7 and PSA-YOLOv7.

Table 4
Performance comparison of the SOTA models.

Model	mAP	FPS	Params
Faster RPAN-CNN [15]	73.6	6.3	260.5 M
SSD [41]	79.9	15.8	23.8 M
BAF-Detector [8]	80.6	7.6	120.9 M
EfficientDet-D0 [42]	78.7	83.3	3.9 M
CornerNet [43]	77.3	6.6	200.9 M
YOLOv5s [44]	81.9	111.1	7.1 M
YOLOv7 [14]	82.0	115	36.9 M
YOLOv8s [45]	82.2	121	21.5 M
PSA-YOLOv7 (ours)	**82.9**	114.5	37.5 M

attention outside the bounding box. This further explains why YOLOv7 performs worse in anomaly detection compared to PSA-YOLOv7.

4.4. Comparison of the PSA-YOLOv7 with SOTA models

In this section, we compare the performance of our proposed PSA-YOLOv7 framework with existing methods for PV cell anomaly detection. We present a comprehensive analysis of the experimental results, discussing the detection accuracy, and computational efficiency of the different approaches.

As shown in Table 4, the detection accuracy of the proposed framework and the existing methods was assessed using mAP. Our proposed PSA-YOLOv7 consistently outperformed the traditional computer vision-based techniques and other CNN-based approaches in terms of all evaluation metrics, demonstrating its superior detection capabilities. The improved accuracy can be attributed to the integration of Partial Convolution and Switchable Atrous Convolution, which enable the model to better handle irregular defects, missing data, and varying defect scales in the PV cell images.

In comparison with the previous YOLO family models, our proposed framework also achieved better detection accuracy, indicating the effectiveness of the architectural improvements and the incorporation of Partial Convolution and Switchable Atrous Convolution.

One of the key advantages of the YOLO family models [14,37,44] is their real-time object detection capabilities, which are essential for large-scale PV cell monitoring systems. In our experiments, the proposed PSA-YOLOv7 maintained a high detection speed, comparable to that of the previous YOLO models. This result demonstrates that our proposed framework is not only more accurate but also computationally efficient, making it suitable for real-time PV cell anomaly detection.

In comparison with traditional computer vision-based techniques and other CNN-based approaches, the proposed framework exhibited superior computational efficiency, further highlighting its suitability for large-scale PV cell monitoring applications.

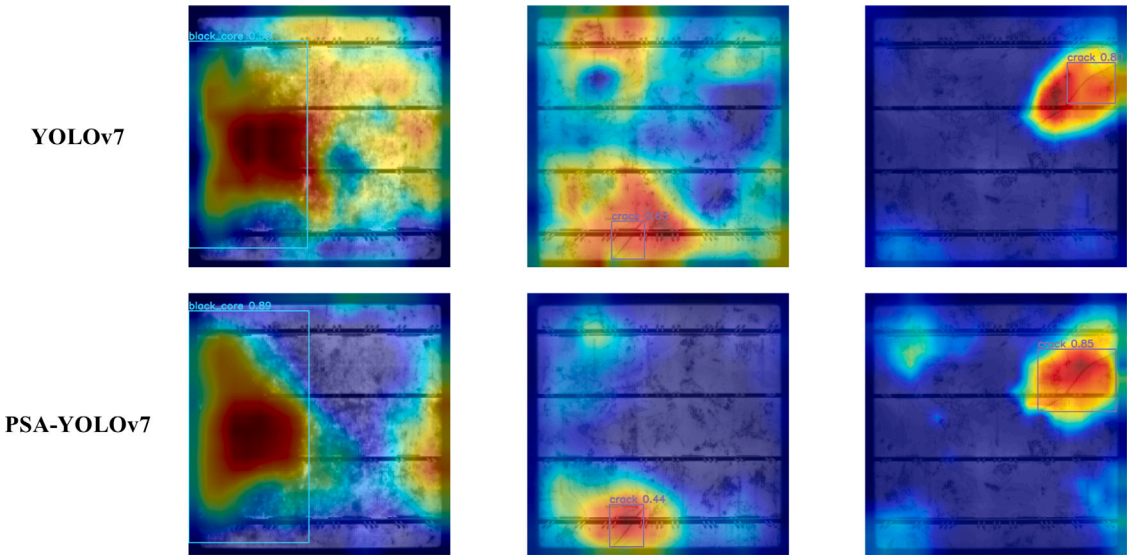


Fig. 10. The attention map visualization of YOLOv7 and PSA-YOLOv7.

Table 5
Comparison of different positions of PELAN module.

C5	C7	C9	C11	Params (M)	mAP@0.5 (%)
✓				37.1	82.1
✓	✓			36.6	82.2
✓	✓	✓		34.6	**82.5**
✓	✓	✓	✓	33.4	82.2

Table 6
Comparison of different positions of SAELAN module.

C5	C7	C9	C11	Params (M)	mAP@0.5 (%)
✓				37.4	82.1
	✓			36.8	82.0
		✓		35.7	82.4
			✓	37.6	**82.9**

4.5. Ablation study

In this section, we validate the designs of the proposed PSA-YOLOv7 for PV cell anomaly detection by experiment. We present an in-depth analysis of the contributions of each component to the overall anomaly detection performance and computational efficiency of the framework. Similar to the open sourced code of YOLOv7, we named the first module of the YOLOv7 backbone as C1, the second module as C2, and so on. The four ELAN modules of the YOLOv7 backbone are C5, C7, C9, and C11. Since both PELAN and SAELAN proposed in this paper replace these parts, the experiments in the ablation study part are mainly carried out on the C5, C7, C9, and C11.

4.5.1. Impact of PELAN module

The PELAN module plays a significant role in enhancing the detection performance of the proposed framework, particularly for PV cell images with irregular defects. By automatically adapting to the irregularities in the input images, Partial Convolution enables more accurate and robust feature extraction, resulting in improved anomaly detection accuracy.

As shown in Table 5, the experimental results show that the PELAN module leads to notable improvements in mAP compared to the baseline YOLOv7 model. These results confirm the effectiveness of Partial Convolution in addressing the challenges associated with irregular defects in PV cell images.

4.5.2. Impact of SAELAN module

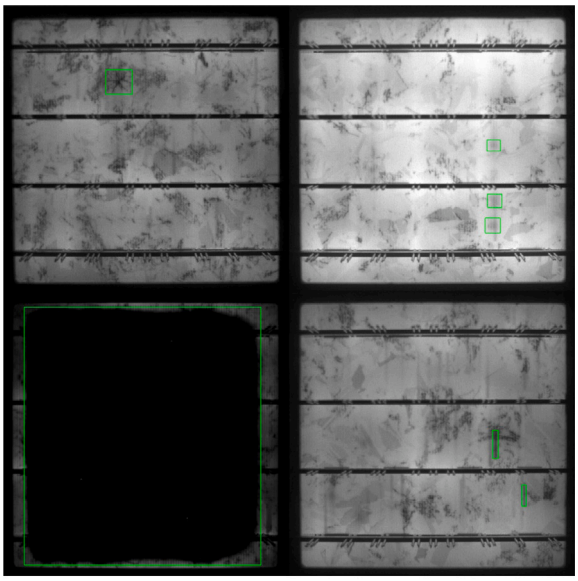
The SAELAN module contributes to the enhanced detection performance of the proposed framework by providing a more flexible and adaptive convolution operation that can handle a wide range of defect scales. By adjusting its receptive field size based on the input data, SAConv enables the model to effectively capture defects and objects of varying sizes in the PV cell images. In this subsection of the experiment, our baseline model consists of PELAN modules for C5, C7, and C9, and an ELAN module for C11. The purpose of this experiment is to replace these modules with SAELAN to investigate the best position for SAELAN to achieve the best performance.

As shown in Table 6, the experimental results indicate that the SAELAN module that integration of Switchable Atrous Convolution results in significant improvements in detection accuracy, particularly for defects that span a wide range of scales. The improved performance can be observed through the increased mAP compared to the baseline YOLOv7 model, highlighting the effectiveness of SAConv in detecting defects across different scales. The integration of Partial Convolution and Switchable Atrous Convolution does not significantly affect the computational complexity of the YOLOv7 architecture, allowing it to maintain its real-time detection capabilities. Our experiments show that the proposed framework maintains a high detection speed, comparable to that of the baseline YOLOv7 model, demonstrating its suitability for real-time PV cell anomaly detection.

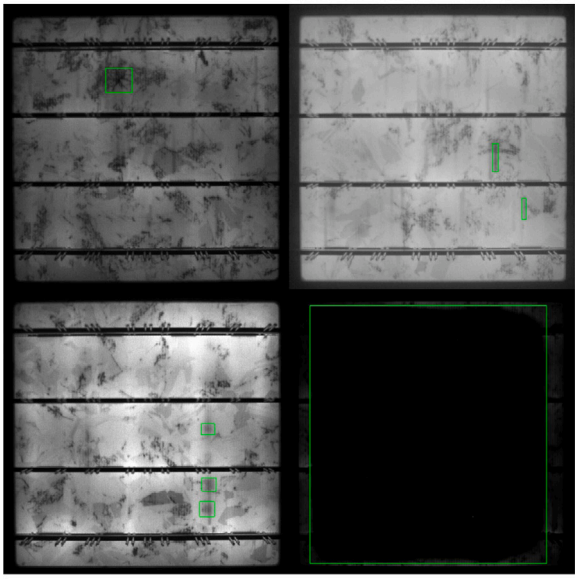
4.5.3. Impact of data augmentations

To improve the performance of our proposed PSA-YOLOv7 model, we have used data augmentation techniques to create additional training data. Data augmentation is a process of generating new training images from the existing dataset by applying transformations such as flipping, rotating, and scaling to the original images.

In this section, we compare the Mosaic data augmentation introduced in the YOLOv4 and RMosaic data augmentation proposed in this work. The Mosaic data augmentation combines four training images into a single image, while RMosaic randomly selects four contrast-augmented training images to create a mosaic image. The RMosaic method is introduced to overcome the limitations of the Mosaic method that may occur in a real production environment. As shown in Fig. 11, the training sample of Mosaic and RMosaic is presented. The main



(a) Mosaic augmentation



(b) RMosaic augmentation

Fig. 11. Visualization of Mosaic augmentation and the proposed RMosaic augmentation.

Table 7
Results of PSA-YOLOv7 with and without data augmentations used in this paper.

Data augmentations	mAP@0.5 (%)
✓	**82.9**
	82.5

difference between RMosaic and Mosaic is the adoption of random contrast enhancement for each sub-image.

To evaluate the effectiveness of RMosaic data augmentation, we have compared it with the original Mosaic data augmentation in terms of mAP. Table 7 presents the comparison results, the mAP value is higher when using RMosaic data augmentation, indicating that the proposed method is more effective in detecting anomalies.

Meanwhile, to verify the robustness of RMosaic in real industrial environments, following the setup in Ref. [46], we selected common corruptions in cameras such as Gaussian Noise, Uniform Noise, and Motion Blur for experimentation. Robustness to various types of noise, motions, and varying lighting conditions is crucial for PV cell anomaly detection systems. As shown in Fig. 12, the results show that the RMosaic data augmentation method has a clear advantage compared to the original Mosaic method under various common corruptions. This is because the RMosaic technique avoids selecting four training images that simulate corruption in real industrial environments, which can adversely impact the model's training performance.

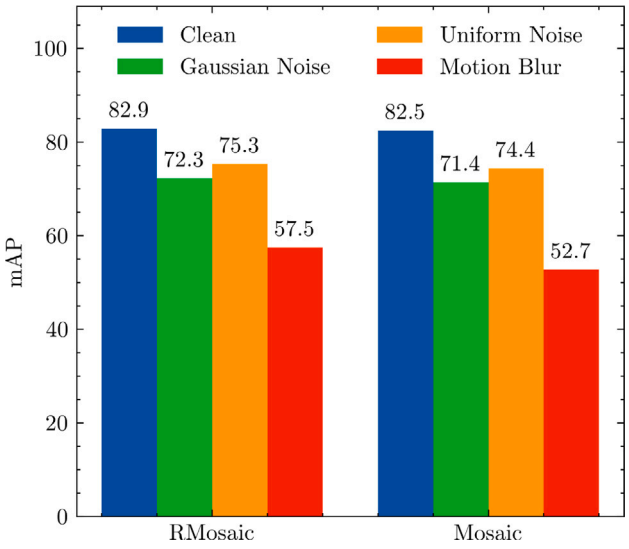


Fig. 12. The robustness of RMosaic and Mosaic data augmentations.

5. Discussion

In this section, we discuss the implications of our proposed PSA-YOLOv7 framework for PV cell anomaly detection, its potential applications in solar energy systems, and the limitations and future research directions.

5.1. Implications of the proposed PSA-YOLOv7 for PV cell anomaly detection

The proposed PSA-YOLOv7 framework, offers several implications for PV cell anomaly detection:

Improved detection accuracy: The integration of Partial Convolution and Switchable Atrous Convolution significantly enhances the detection accuracy of the framework, particularly for images with irregular defects and defects of varying sizes. This improvement can lead to more reliable and efficient PV cell monitoring and maintenance.

Real-time detection capabilities: Despite the integration of advanced techniques, the proposed framework maintains real-time detection capabilities, which are essential for large-scale PV cell monitoring systems. This feature enables the framework to be readily applied in real-world solar energy installations, ensuring timely detection and remediation of defects.

Robustness to various defects and conditions: The proposed framework demonstrates robustness in handling various defect types, occlusions, and varying lighting conditions. This robustness ensures that the system can effectively detect anomalies across diverse PV cell installations, enhancing the overall reliability of solar energy systems.

5.2. Potential applications in solar energy systems

The proposed PSA-YOLOv7 framework for PV cell anomaly detection can be applied in various solar energy systems to ensure efficient operation, such as quality control in PV cell manufacturing. The

proposed framework can be used for quality control during PV cell manufacturing processes, detecting defects and irregularities in the produced PV cells before they are installed in solar energy systems.

5.3. Limitations and future research directions

Despite the promising results of our proposed PSA-YOLOv7 framework for PV cell anomaly detection, there are several limitations and potential areas for future research:

Performance under extreme conditions: The performance of the proposed framework under extreme weather conditions, such as heavy snow or sandstorms, remains to be evaluated. Future research could focus on developing more robust models that can withstand these extreme conditions, ensuring reliable anomaly detection in various environments.

Integration with other sensors and data sources: The proposed framework solely relies on PV cell images for anomaly detection. Integrating additional sensors and data sources, such as temperature, LiDAR [47] and humidity sensors, could provide valuable contextual information and further improve detection accuracy.

Handling new and rare defect types: While the proposed framework demonstrates robustness in handling various defect types, its performance in detecting new and rare defect types remains to be explored. Future research could investigate methods for effectively detecting and classifying these rare defects, enhancing the overall reliability of PV cell monitoring systems.

6. Conclusion

In this paper, we have presented a novel PSA-YOLOv7 framework for fast anomaly detection of photovoltaic (PV) cells. We incorporate advanced techniques such as Partial Convolution and Switchable Atrous Convolution to address the challenges associated with irregular defects and defects of varying sizes. A comprehensive evaluation of the proposed framework, demonstrates its superior performance in terms of detection accuracy, computational efficiency, and robustness compared to traditional computer vision-based techniques, other CNN-based approaches, and previous YOLO family models. A thorough analysis of the impacts of Partial Convolution and Switchable Atrous Convolution on the detection performance, and computational efficiency of the proposed framework, highlighting their contributions to the overall effectiveness of the system.

The proposed PSA-YOLOv7-based framework holds significant implications for PV cell anomaly detection, as it offers improved detection accuracy, real-time detection capabilities, and robustness to various defects and conditions. These advancements can lead to more reliable and efficient PV cell monitoring and maintenance, ultimately contributing to the optimization of solar energy systems and their overall performance. The proposed framework is applicable to a wide range of solar energy installations, from large-scale solar farms to rooftop solar systems.

The success of our proposed PSA-YOLOv7 framework demonstrates the potential of deep learning-based anomaly detection in PV cells. As the demand for renewable energy sources continues to grow, so does the need for efficient and reliable monitoring systems to ensure the optimal performance of solar energy installations. We anticipate that deep learning-based approaches, such as the one presented in this paper, will play an increasingly important role in the development of advanced PV cell monitoring systems. Future research efforts should focus on addressing the limitations of current methods and exploring new techniques to further improve detection accuracy, computational efficiency, and robustness under various conditions. By continuing to advance deep learning-based anomaly detection in PV cells, researchers and practitioners can contribute to the growth and sustainability of solar energy systems, ultimately supporting the global transition toward cleaner and more reliable energy sources.

CRediT authorship contribution statement

Jinlai Zhang: Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization. **Wenjie Yang:** Project administration, Methodology, Funding acquisition. **Yumei Chen:** Software, Resources. **Mingkang Ding:** Software, Resources. **Huiling Huang:** Visualization, Validation. **Bingkun Wang:** Visualization, Validation. **Kai Gao:** Investigation, Data curation. **Shuhan Chen:** Investigation, Data curation. **Ronghua Du:** Investigation, Funding acquisition.

Declaration of competing interest

The authors declare that there is no conflict of interest in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The project was supported by the National Natural Science Foundation of China (grant number 61973047) and Scientific Research Fund of Hunan Provincial Education Department, China (Grant number 23B0884).

References

[1] Khan W, Walker S, Zeiler W. Improved solar photovoltaic energy generation forecast using deep learning-based ensemble stacking approach. Energy 2022;240:122812.
[2] Ren H, Xu C, Ma Z, Sun Y. A novel 3D-geographic information system and deep learning integrated approach for high-accuracy building rooftop solar energy potential characterization of high-density cities. Appl Energy 2022;306:117985.
[3] Hernandez RR, Armstrong A, Burney J, Ryan G, Moore-O'Leary K, Diédhiou I, et al. Techno–ecological synergies of solar energy for global sustainability. Nat Sustain 2019;2(7):560–8.
[4] Liu W, Liu Y, Yang Z, Xu C, Li X, Huang S, et al. Flexible solar cells based on foldable silicon wafers with blunted edges. Nature 2023;617(7962):717–23.
[5] Ghadikolaei SSC. An enviroeconomic review of the solar PV cells cooling technology effect on the CO2 emission reduction. Sol Energy 2021;216:468–92.
[6] Wang H, Lei Z, Zhang X, Zhou B, Peng J. A review of deep learning for renewable energy forecasting. Energy Convers Manage 2019;198:111799.
[7] Nam K, Hwangbo S, Yoo C. A deep learning-based forecasting model for renewable energy scenarios to guide sustainable energy policy: A case study of Korea. Renew Sustain Energy Rev 2020;122:109725.
[8] Su B, Chen H, Zhou Z. BAF-detector: An efficient CNN-based detector for photovoltaic cell defect detection. IEEE Trans Ind Electron 2021;69(3):3161–71.
[9] Su B, Zhou Z, Chen H. PVEL-AD: A large-scale open-world dataset for photovoltaic cell anomaly detection. IEEE Trans Ind Inf 2022;19(1):404–13.
[10] Dino IG, Sari AE, Iseri OK, Akin S, Kalfaoglu E, Erdogan B, et al. Image-based construction of building energy models using computer vision. Autom Constr 2020;116:103231.
[11] Xie J, Yang R, Gooi HB, Nguyen HD. PID-based CNN-LSTM for accuracy-boosted virtual sensor in battery thermal management system. Appl Energy 2023;331:120424.
[12] Niu D, Yu M, Sun L, Gao T, Wang K. Short-term multi-energy load forecasting for integrated energy systems based on CNN-BiGRU optimized by attention mechanism. Appl Energy 2022;313:118801.
[13] Jocher G, Stoken A, Borovec J, Chaurasia A, Changyu L, Laughing A, et al. Ultralytics/yolov5: v5. 0-YOLOv5-P6 1280 models AWS supervise. ly and YouTube integrations. 2021, Zenodo, 11.
[14] Wang C-Y, Bochkovskiy A, Liao H-YM. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: CVPR. 2023.
[15] Su B, Chen H, Chen P, Bian G, Liu K, Liu W. Deep learning-based solar-cell manufacturing defect detection with complementary attention network. IEEE Trans Ind Inf 2020;17(6):4084–95.
[16] Harrou F, Taghezouit B, Bouyeddou B, Sun Y, Arab AH. Fault detection in solar PV systems using hypothesis testing. In: 2021 IEEE 19th international conference on industrial informatics. IEEE; 2021, p. 1–6.
[17] Vergura S. Hypothesis tests-based analysis for anomaly detection in photovoltaic systems in the absence of environmental parameters. Energies 2018;11(3):485.

[18] Fazai R, Abodayeh K, Mansouri M, Trabelsi M, Nounou H, Nounou M, et al. Machine learning-based statistical testing hypothesis for fault detection in photovoltaic systems. Sol Energy 2019;190:405–13.
[19] Buddha S, Braun H, Krishnan V, Tepedelenlioglu C, Spanias A, Yeider T, et al. Signal processing for photovoltaic applications. In: 2012 IEEE international conference on emerging signal processing applications. IEEE; 2012, p. 115–8.
[20] Harrou F, Taghezouit B, Sun Y. Improved k NN-based monitoring schemes for detecting faults in PV systems. IEEE J Photovolt 2019;9(3):811–21.
[21] Arena E, Corsini A, Ferulano R, Iuvara DA, Miele ES, Ricciardi Celsi L, et al. Anomaly detection in photovoltaic production factories via Monte Carlo pre-processed principal component analysis. Energies 2021;14(13):3951.
[22] Tan H, Guo Z, Zhang H, Chen Q, Lin Z, Chen Y, et al. Enhancing PV panel segmentation in remote sensing images with constraint refinement modules. Appl Energy 2023;350:121757.
[23] Buerhop-Lutz C, Deitsch S, Maier A, Gallwitz F, Berger S, Doll B, et al. A benchmark for visual identification of defective solar cells in electroluminescence imagery. In: 35th European PV solar energy conference and exhibition, vol. 12871289, 2018, p. 1287–9.
[24] Deitsch S, Christlein V, Berger S, Buerhop-Lutz C, Maier A, Gallwitz F, et al. Automatic classification of defective photovoltaic module cells in electroluminescence images. Sol Energy 2019;185:455–68.
[25] Li X, Yang Q, Lou Z, Yan W. Deep learning based module defect analysis for large-scale photovoltaic farms. IEEE Trans Energy Convers 2018;34(1):520–9.
[26] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016, p. 779–88.
[27] Cai J, Xu K, Zhu Y, Hu F, Li L. Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest. Appl Energy 2020;262:114566.
[28] Chen J, Kao S-h, He H, Zhuo W, Wen S, Lee C-H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks. 2023, arXiv preprint arXiv:2303.03667.
[29] Qiao S, Chen L-C, Yuille A. Detectors: Detecting objects with recursive feature pyramid and switchable atrous convolution. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021, p. 10213–24.
[30] Shi Z, Yao W, Zeng L, Wen J, Fang J, Ai X, et al. Convolutional neural network-based power system transient stability assessment and instability mode prediction. Appl Energy 2020;263:114586.
[31] Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017, p. 1251–8.
[32] LeCun Y, Bengio Y, Hinton G. Deep learning. nature 2015;521(7553):436–44.
[33] Chen F, Wu F, Xu J, Gao G, Ge Q, Jing X-Y. Adaptive deformable convolutional network. Neurocomputing 2021;453:853–64.
[34] Jumbo O, Moghaddass R. Resource optimization and image processing for vegetation management programs in power distribution networks. Appl Energy 2022;319:119234.
[35] Zhang J, Chen L, Ouyang B, Liu B, Zhu J, Chen Y, et al. Pointcutmix: Regularization strategy for point cloud classification. Neurocomputing 2022.
[36] Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: An imperative style, high-performance deep learning library. In: Advances in neural information processing systems, vol. 32, 2019.
[37] Yang W, Wu J, Zhang J, Gao K, Du R, Wu Z, et al. Deformable convolution and coordinate attention for fast cattle detection. Comput Electron Agric 2023;211:108006.
[38] Zhang J, Meng Y, Wu J, Qin J, Yao T, Yu S, et al. Monitoring sugar crystallization with deep neural networks. J Food Eng 2020;280:109965.
[39] Zhang Q, Yang L, Guo W, Qiang J, Peng C, Li Q, et al. A deep learning method for lithium-ion battery remaining useful life prediction based on sparse segment data via cloud computing system. Energy 2022;241:122716.
[40] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision. 2017, p. 618–26.
[41] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y, et al. Ssd: Single shot multibox detector. In: Computer vision–ECCV 2016: 14th European conference, Amsterdam, the Netherlands, October 11–14, 2016, proceedings, Part I 14. Springer; 2016, p. 21–37.
[42] Cai Z, Vasconcelos N. Cascade R-CNN: High quality object detection and instance segmentation. IEEE Trans Pattern Anal Mach Intell 2019;1. http://dx.doi.org/10.1109/tpami.2019.2956516.
[43] Law H, Deng J. Cornernet: Detecting objects as paired keypoints. In: 15th European conference on computer vision. Springer Verlag; 2018, p. 765–81.
[44] Glenn J. YOLOv5 release v6.1. 2022, https://github.com/ultralytics/yolov5/releases/tag/v6.1.
[45] Glenn J. Ultralytics YOLOv8. 2023, https://github.com/ultralytics/ultralytics.
[46] Dong Y, Kang C, Zhang J, Zhu Z, Wang Y, Yang X, et al. Benchmarking robustness of 3D object detection to common corruptions. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023, p. 1022–32.
[47] Han B, Wei J, Zhang J, Meng Y, Dong Z, Liu H. GardenMap: Static point cloud mapping for garden environment. Comput Electron Agric 2023;204:107548.