

Article

GBH-YOLOv5: Ghost Convolution with BottleneckCSP and Tiny Target Prediction Head Incorporating YOLOv5 for PV Panel Defect Detection

Longlong Li ^{1,2}, Zhifeng Wang ^{1,*}  and Tingting Zhang ²¹ Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan 430079, China² School of Computer Science and Engineering, Northwest Normal University, Lanzhou 730070, China

* Correspondence: zfwang@ccnu.edu.cn

Abstract: Photovoltaic (PV) panel surface-defect detection technology is crucial for the PV industry to perform smart maintenance. Using computer vision technology to detect PV panel surface defects can ensure better accuracy while reducing the workload of traditional worker field inspections. However, multiple tiny defects on the PV panel surface and the high similarity between different defects make it challenging to accurately identify and detect such defects. This paper proposes an approach named Ghost convolution with BottleneckCSP and a tiny target prediction head incorporating YOLOv5 (GBH-YOLOv5) for PV panel defect detection. To ensure better accuracy on multiscale targets, the BottleneckCSP module is introduced to add a prediction head for tiny target detection to alleviate tiny defect misses, using Ghost convolution to improve the model inference speed and reduce the number of parameters. First, the original image is compressed and cropped to enlarge the defect size physically. Then, the processed images are input into GBH-YOLOv5, and the depth features are extracted through network processing based on Ghost convolution, the application of the BottleneckCSP module, and the prediction head of tiny targets. Finally, the extracted features are classified by a Feature Pyramid Network (FPN) and a Path Aggregation Network (PAN) structure. Meanwhile, we compare our method with state-of-the-art methods to verify the effectiveness of the proposed method. The proposed PV panel surface-defect detection network improves the mAP performance by at least 27.8%.



Citation: Li, L.; Wang, Z.; Zhang, T. GBH-YOLOv5: Ghost Convolution with BottleneckCSP and Tiny Target Prediction Head Incorporating YOLOv5 for PV Panel Defect Detection. *Electronics* **2023**, *12*, 561. <https://doi.org/10.3390/electronics12030561>

Academic Editor: Valentina E. Balas

Received: 24 December 2022

Revised: 17 January 2023

Accepted: 19 January 2023

Published: 21 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: YOLOv5; PV defect detection; BottleneckCSP; GhostConv; tiny target prediction head

1. Introduction

The high public demand for attention to environmental issues has become an essential indicator in the energy sector; to deal with the ecological problems that affect all human beings, reduce greenhouse gas emissions, and avoid the catastrophic consequences of climate change, countries around the world and the World Health Organization have developed relevant policies [1]. The terms “carbon peak” and “carbon neutral” are the most critical energy and environmental policies to cope with the global warming problem. To achieve the “double carbon” goal, renewable energy, represented by photovoltaic power generation, is undoubtedly the main force [2]. As an essential part of the development of the PV industry, the fault detection of PV panels is of great significance in promoting the development of PV energy [3]. With the development of artificial intelligence, the intelligent detection of PV panel faults is becoming a feasible and promising solution. Using machine vision techniques to identify surface defects in PV panels has become an essential technical basis for building intelligent PV inspection systems [4,5]. Deep learning techniques can significantly improve detection efficiency, provide solutions for the competent inspection of PV power plants, and guide power plants’ operation and maintenance procedures [6,7].

The current processing techniques for PV panel images are mainly divided into two categories [8]. The first category is the traditional machine learning methods, which mostly

rely on manually designed extractors and require the manual construction of complex recognition relationships [9], and their generalization ability and robustness could be better [10–12]. YOLO and Region-CNN (R-CNN) algorithms, represented by deep learning techniques, are another class of methods that rely mainly on learning a large number of samples to obtain a deep dataset feature representation with better generalization ability and robustness [13,14]. Inspired by the previous research, we use YOLOv5 as the primary network framework, which is fast while maintaining good accuracy [15,16].

This paper also introduces the BottleneckCSP module and the Ghost convolution mechanism, which help the model to obtain more information about the characteristics while maintaining the detection speed. YOLOv5s are used to detect five types of defects on the surface of PV panels: broken, hot_spot, black_border, scratch, and no_electricity. At the same time, this paper compares five detection frameworks within the same family as YOLOv3: the bipartite target detection methods Faster-RCNN and Mask-RCNN, the traditional machine learning method SVM, and Single Shot MultiBox Detector. The contributions of this paper can be summarized as follows:

- To the best of our knowledge, we are the first to apply the YOLOv5 structure to tackle the task of detecting defects on PV panels. This study utilizes the fast inference speed and high detection accuracy of YOLOv5 to obtain a combination of detection speed and accuracy on the PV Multi-Defect dataset, which enables accurate and rapid detection of various types of defects in PV panels and significantly reduces the missed detection of minor defects.
- According to the PV panel defect detection task, the structure of YOLOv5 is improved and innovated in this paper. Firstly, the semantic depth information of PV panel images is obtained using the BottleneckCSP module, improving detection accuracy. Secondly, the added detection head for tiny targets alleviates the negative impact of drastic scale changes and improves the small target misdetection phenomenon. On this basis, Ghost convolution is introduced instead of conventional convolution, and we call this structure GBH-YOLOv5, which can perform the PV panel defect detection task well. The implementation codes of this research are released at <https://github.com/CCNUZFW/GBH-YOLOv5> (accessed on 23 December 2022).
- In this paper, a new database dedicated to PV defect detection is constructed, which includes 5 types of defect targets and 1108 images with an image size of 600×600 pixels. There are 886 images in the training set and 222 in the validation set. Moreover, the database is publicly released to promote the field at the following links: <https://github.com/CCNUZFW/PV-Multi-Defect> (accessed on 23 December 2022).
- By comparing this method with five state-of-the-art methods, the proposed PV panel surface defect approach has improved the mAP by at least 27.8%, and the single image detection time consumed is in the same order of magnitude, balancing detection accuracy and detection speed. It provides significant advantages in identifying various types of defects on the surface of PV panels.

The remainder of the paper is structured as follows: Section 2 describes PV panel defect detection and the related studies on YOLO. Section 3 describes the defect detection process and the network framework, and in Section 4, comparison and ablation experiments are performed. Finally, the conclusions of the article are stated in Section 5.

2. Related Work

This section presents two parts of the related work: (2.1) the current state of research on PV panel defect detection and (2.2) the development of target detection based on the YOLO algorithm.

2.1. PV Panel Defect Detection

With the progress in energy structures, photovoltaic power generation, considered the most promising approach, is developing rapidly and playing a significant role in energy security, national income, public health, and environmental protection. As an essential

component of a PV power generation system, PV panels are subject to challenging working environments and prone to faults, which affect the operation and lifetime of the entire PV system. Therefore, the fault detection of PV panels is the key to improving PV systems' efficiency, reliability, and lifecycle. There are three mainstream detection methods: image processing-based methods, electrical detection-based methods, and machine learning-based methods.

(1). Image processing-based methods: Among the image processing-based methods, various imaging solutions exist depending on the different characteristics of the panels. In thermal imaging, an infrared camera is used to scan the PV array, which is suitable for inspecting large PV plants. The ultrasonic imaging inspection method is used primarily for detecting cracks before the production of PV modules; an electroluminescence imaging (EL image) solution is a unique image presented by the panel at a specific voltage, which is more expensive to detect. In conclusion, imaging solutions rely on the various types of image features produced by PV panels under different techniques to determine their fault type. In [17], by varying the modulation of the injected current, the panel image was made to exhibit certain features that allowed the detection of different types of shunt faults. In [18], the authors have verified that high accuracy fault identification is possible by performing thermal imaging analysis of PV panels and using radiation sensors. V. Kirubakaran et al. [19] use a thermal imaging system combined with image processing to record PV panel failure points.

(2). Electrical detection-based methods: Electrical detection-based methods include basic current–voltage measurement techniques, advanced Climate-Independent Data (CID) plans, and power loss detection methods that enable fault detection and classification. Electrical detection methods diagnose specific faults based on different electrical characteristics. In [20], the Time Domain Reflector (TDR) technique is used to locate PV module faults based on the delay between the injected and reflected signals. In addition, the output voltage and current of the PV panel string are measured to identify possible faults in advance. A.L. et al. [21] constructed a model for local defect and thermal breakdown detection of PV panels based on thermal images and IV curves.

(3). Machine learning-based methods: Since the performance and efficiency of PV cells are subject to various conditions, many problems are difficult to define in specific projects. Machine learning techniques can overcome these difficulties very well due to their self-learning nature, making them widely used in this type of detection [22]. In [23], the authors used grayscale cogeneration matrices to extract image features generated by infrared imaging techniques to monitor defects in panel modules. In [13], the authors used support vector machines (RBF kernel) and random forest algorithms to construct detection models to obtain the desired detection accuracy in the electroluminescence dataset. Jumaboev et al. verified the feasibility of deep learning techniques in photovoltaic inspection by using several deep learning models [24]. F.L. et al. [25] proposed a semi-supervised anomaly detection model based on adversarial generative networks for PV panel defect detection. In [26], an automatic detection method for optoelectronic components was proposed based on texture analysis and supervised learning for the processing of infrared images. Chiwu Bu et al. [27] used LDA and QDA supervised learning algorithms for the processing and defect identification of photovoltaic panel thermographic sequences.

2.2. Target Detection Based on YOLO

The YOLO algorithm is a one-stage target detection method proposed by Joseph Redmon. He converted the object detection problem into a regression problem by discarding the branching phase of candidate box extraction in the two-stage target detection algorithm and completing the determination of the entire category and the regression of the position in a single network [28].

In YOLOv2, the authors introduced anchor boxes and batch normalization to improve the problem of the low detection accuracy of the v1 model [29]. YOLOv3 built a new Darknet with 53 residual networks based on YOLOv2 and passed feature pyramid networks

for multiscale fusion prediction, which improved the detection accuracy of small and heavy targets [30]. YOLOv4 constructed a simple target detection model, which reduced the training threshold of this algorithm [31]. YOLOv5 constructed five models of YOLOv5, N/S/M/L/X, based on the scaling of different channels and the model size [32].

In the second half of 2022, YOLOv6 and YOLOv7 were released almost simultaneously. The Meituan technical team introduced the RepVGG structure in YOLOv6, which was more adaptable to GPU devices and simplified the adaptation work during engineering deployments [33]. YOLOv7 used module re-referencing and a dynamic label assignment strategy, which made it faster and more accurate in the range of 5 FPS to 160 FPS and exceeded known target detectors [34].

In the PV panel surface-defect detection system used in this paper, the C3 module in the network structure of YOLOv5s is replaced with the BottleneckCSP module, and the number of detection heads is increased so that the grid acquires more feature information and improves the detection capability for small targets. In addition, the conventional convolution is replaced with Ghost convolution to reduce the model's parameters and the computational effort in the inference process.

3. Methods

The task of finding defects in PV panels has two characteristics: first, it must precisely pinpoint the fault; second, it must accurately define the defect attributes. The models in the YOLOv5 series are quite good at localization, and YOLOv5s is also lighter. Compared to other deep learning techniques, YOLOv5s is able to strike a balance between detection accuracy and speed, making it a good choice for this purpose.

The network structure of this research is implemented by improving the network structure of YOLOv5s, named GBH-YOLOv5, as shown in Figure 1.

The network consisted of four parts: the head, backbone, neck, and prediction. The head was the input, and the Mosaic data enhancement and adaptive image scaling were used to expand the data samples, enrich the test data set, and increase the robustness of the network on the one hand; on the other hand, the problem of the uneven distribution of the small target dataset was solved. By setting different anchor frames and constantly updating the difference between the prediction frame and the labeled frame, the adaptive anchor frame updated the network parameters and independently calculated the optimal anchor frame value to learn the feature information about the target better. The backbone network was implemented by Focus, SPP, and BottleneckCSP, based on Ghost convolution, and the input 960×960 pixel image was sliced to obtain a $480 \times 480 \times 12$ feature map; then, a $480 \times 480 \times 32$ feature map was obtained after convolution. The PV panel surface-defect detection network used a Focus structure to reduce the number of network layers and parameters, improve the forward and reverse computation speed, and ensure no information loss in downsampling. In order to extract more feature information, this network used a BottleneckCSP module with two deep convolutions. To reduce the redundant feature information brought by the BottleneckCSP module, Ghost convolution was used instead of the traditional convolution to reduce the parameters and the computation of the model inference process. The prediction was the output side, where CloU Loss was used as the loss function of the bounding box to solve the problem of the non-overlap between the labeled and predicted boxes; the Non-Maximum Suppression (NMS) mechanism was used to enhance the recognition ability of multiple targets and fuzzy targets. In order to ensure that the scratch and defective break were accurately recognized, a detection head was added to the output of the network to realize the detection of tiny targets.

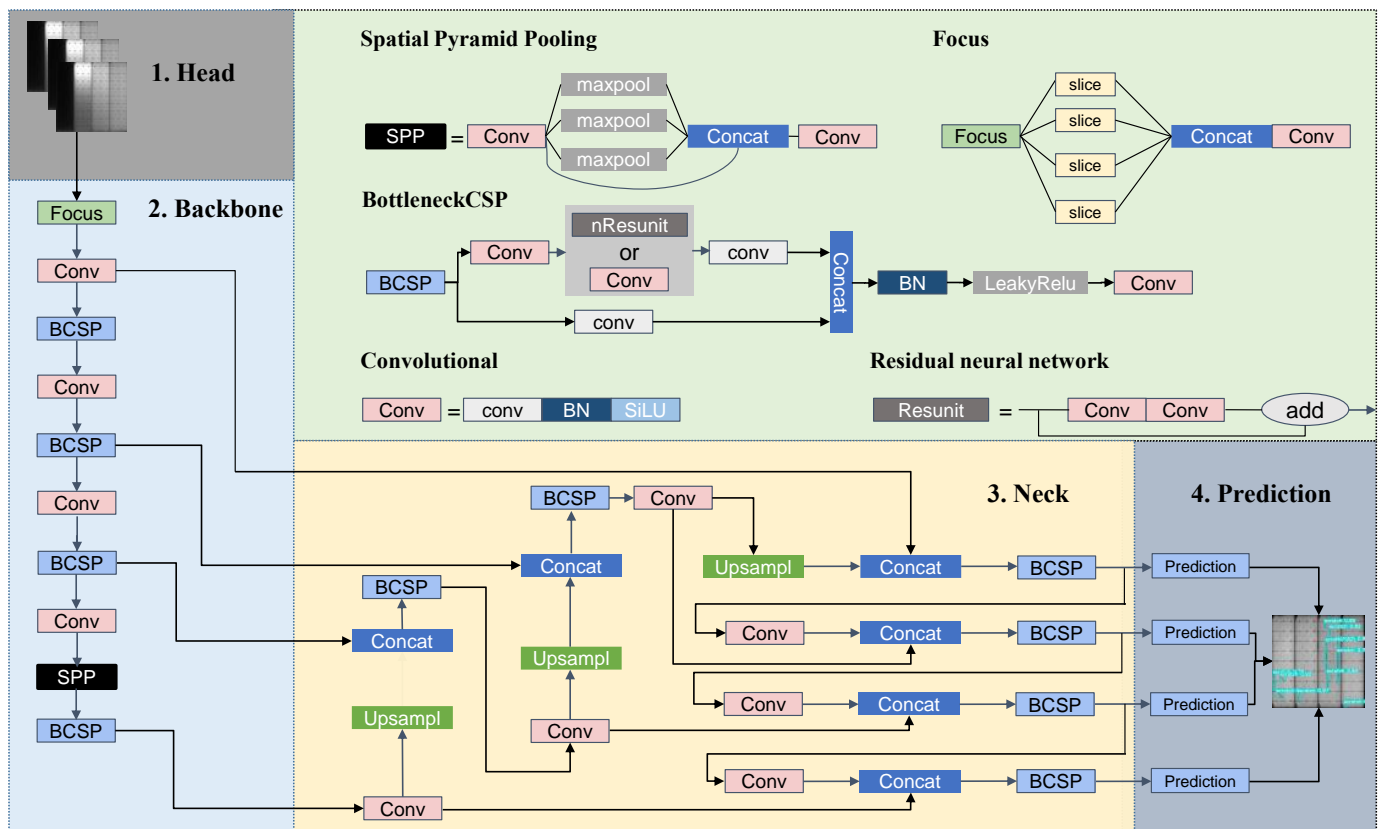


Figure 1. The framework of PV panel surface-defect detection network. Using the BottleneckCSP module in the backbone network and neck network ensures that deeper semantic information of PV panels can be extracted, and a tiny target detection head is added at the prediction end so that the PV panel missed detection phenomenon is improved. Replacing the traditional convolution with Ghost convolution can guarantee the accuracy of PV panel detection without losing speed.

When performing detection, the network divides the input image into $S \times S$ grids (grid detection method) [28]. If the target's center point is inside a grid, that grid is responsible for the object detection grid. The detection idea is shown in Figure 2. When the grid detects the object, it outputs a bounding box, and each bounding box consists of five parameters, i.e., four coordinate parameters and one confidence parameter. t_x and t_y denote the coordinates of the center point of the bounding box, and t_w and t_h represent the width and height of the bounding box, respectively. The confidence level indicates whether the current bounding box includes the object to be detected and its accuracy [35,36].

3.1. PV Panel Defect Detection Process

The detection method proposed in this paper was composed of three processing modules, mainly used for surface-defect detection on the PV panels, as shown in Figure 3.

(1) Input module: This module input the captured images into the PV panel defect detector, which has no requirement for the input size of the images.

(2) PV panel defect detector module: First, the size of the input image was checked, and for images whose size was not 600×600 pixels, a cropping and compression process was performed, which physically enlarged the panel defect and reduced the negative sample information. Then, pretrained weights based on the COCO dataset were used to train in the modified YOLO network.

(3) Output module: Defect detection was performed on the resulting images using the YOLOv5 network.

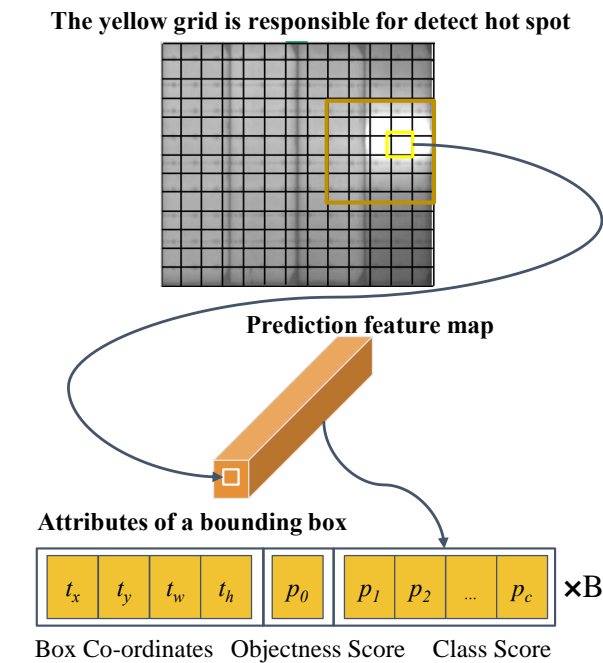


Figure 2. Basic detection idea.

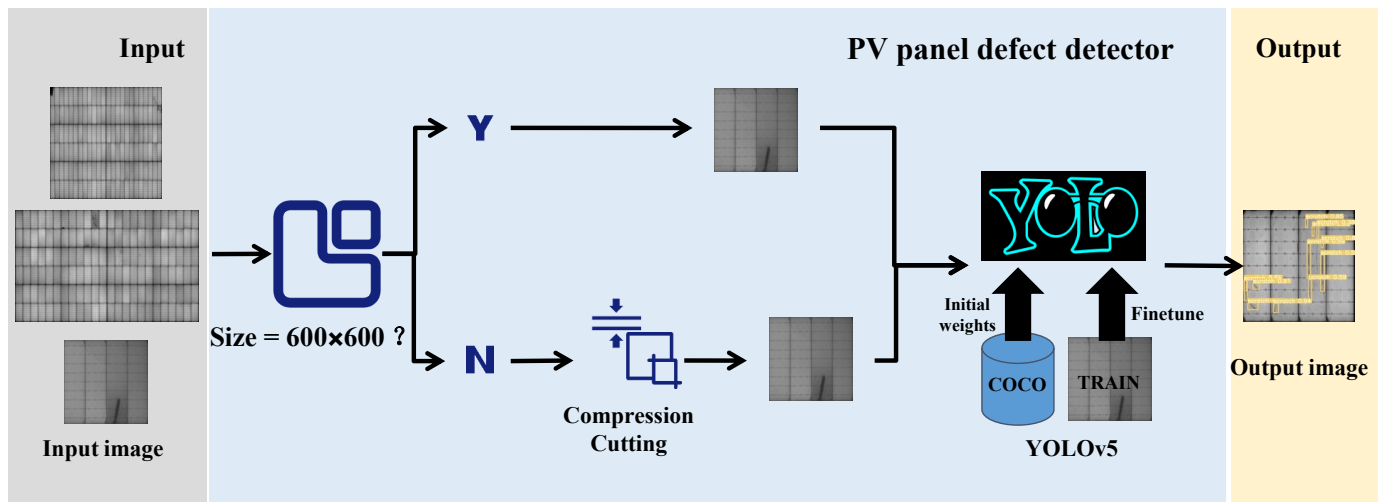


Figure 3. PV panel defect detection process.

3.2. BottleneckCSP Module

The BottleneckCSP was mainly used to extract deep semantic information from images and fuse the feature maps of different scales to enrich the semantic information. The role of its primary structure, bottleneck, combined shallow-level feature maps with deep-level feature maps downward by summation to ensure that the detector maintained good accuracy on multiscale targets. After mixing the CSP, the feature maps were integrated at the network's beginning and end to reflect the gradients' variability. It can be expressed as Equation (1).

$$\begin{aligned} y &= F(x_0) = x_k \\ &= H_k(x_{k-1}, H_{k-1}(x_{k-2}), H_{k-2}(x_{k-3}), \dots, H_1(x_0), x_0). \end{aligned} \quad (1)$$

H_k is the operator function of the k^{th} layer, which usually consists of a convolution layer and an activation function. A y function was introduced to optimize each H function.

$$y = M(x_0', T(F(x_0'')), \quad (2)$$

where x_0 can be divided into two parts along the channel, the T function truncates the gradient flow, and the M function is used to mix the two parts. We obtained an information-rich feature map to retain and accumulate more features from different sensory fields [37].

3.3. Prediction Head for Tiny Targets

We analyzed the PV Multi-Defect dataset and found a considerable proportion of tiny targets (scratches), so we added another prediction head to detect tiny targets. Combined with the other three prediction heads, the four-head structure sufficiently alleviated the negative impact of the drastic scale transformation and thus mitigated the missed detection phenomenon. We added an anchor frame for small targets, enhanced the features from the second layer of the backbone network, and finally added a prediction head for the second layer.

3.4. GhostConv Module

In Ghost convolution, only part of the feature map generated by conventional convolution was used to avoid the redundancy of the feature map. Then, a simple linear transformation was performed on this part of the feature map to achieve the effect of simulating conventional convolution [38]; the convolution process is shown in Figure 4.

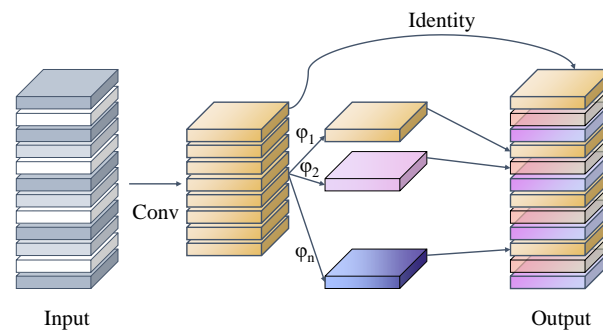


Figure 4. The Ghost convolution process.

In this work, we used $X \in R^{c \times h \times w}$ to represent the input feature map, c to represent the number of channels of the input feature map, and h and w to express the height and width of the feature map, respectively. The conventional convolution is defined as Equation (3).

$$Y = X * f + b. \quad (3)$$

In the above equation, $X \in R^{c \times h' \times w'}$ denotes a feature map with n channels output, and h' and w' denote the height and width of the output feature map, respectively. $*$ denotes the convolution operation, the convolution kernel size is $k * k$, b denotes the bias term, and the regular convolution computation after ignoring the bias term is approximately equal to $h \times w \times c \times n \times w' \times h'$. In the shallow layer of the network, h' and w' are more extensive, and in the deeper layer, n and c are larger. Based on this feature, Ghost convolution was proposed, which consisted of two parts: the regular convolution kernel that outputs a small number of feature maps and the generation of redundant feature maps in a lightweight linear transform layer, which can be expressed as

$$Y' = X * f' + b. \quad (4)$$

The above equation represents a conventional convolutional layer that outputs a small number of feature maps, where $Y' \in R^{h' \times w' \times m}$ represents the output feature and $f' \in R^{c \times k \times k \times m}$ represents the size of this convolutional kernel. The number of channels of the output feature map is smaller than that of the conventional convolutional layer, i.e., $m < n$.

$$y_{ij} = \phi_{ij}(y'_i). \quad (5)$$

Equation (5) denotes the linear transformation layer that represents the generation of redundant feature maps, where y_i denotes the m feature maps of Y' . Each feature map in Y' is subjected to a lightweight linear transformation operation $\phi_{i,j}(j = 1, 2, \dots, s)$ to obtain s feature maps. The last linear transform is forcibly specified as a constant transform if the $d \times d$ convolution is used as the linear transform; so, m feature maps are obtained after linear transformations of $m \times (s - 1)$ feature maps. The total computation using Ghost convolution is $(s - 1) \times m \times h' \times w' \times k \times k$.

4. Experimental Results and Discussion

This section presents the dataset and describes the analysis and preprocessing; then, the target detection field's baseline is discussed, and comparison and ablation experiments are described.

4.1. Dataset for Experiments

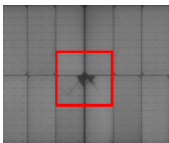
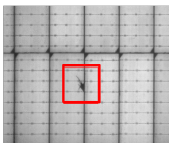
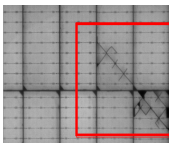
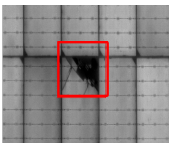
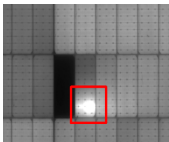
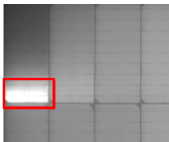
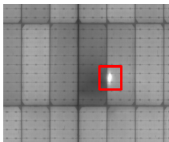
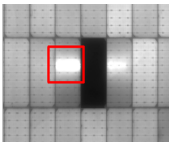
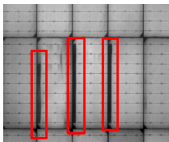
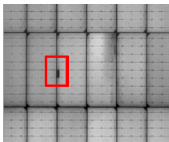
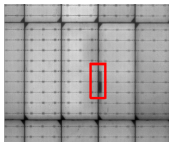
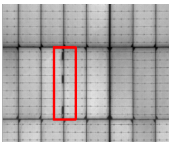
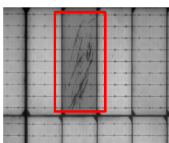
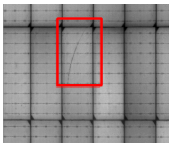
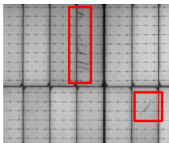
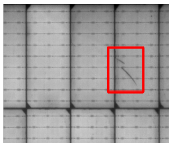
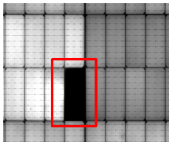
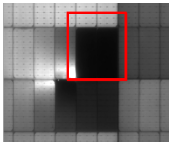
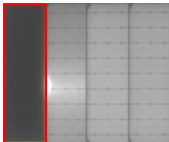
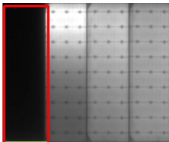
We constructed a publicly available dataset to verify the model's validity and named it the PV Multi-Defect dataset (<https://github.com/CCNUZFW/PV-Multi-Defect> (accessed on 23 December 2022)). The original images for this dataset were taken by the camera from photovoltaic modules with a physical size of 1.65 m \times 0.991 m and a specification of 60 pieces. After grayscale processing, the images were uniformly cropped to 0.491 m \times 0.297 m in accordance with the distribution of defects, and the images that did not demonstrate defects were manually removed. In total, 307 images, each measuring 5800 \times 3504 pixels, were collected, as shown in Figure 5. In this dataset, there were five common defect types, including broken cells, cells with prominent bright spots, cells with regularly shaped black or gray edges, cells with scratches, and cells that were not charged and appeared black. Table 1 shows examples of each defect type. Figure 6 shows the training and validation losses.

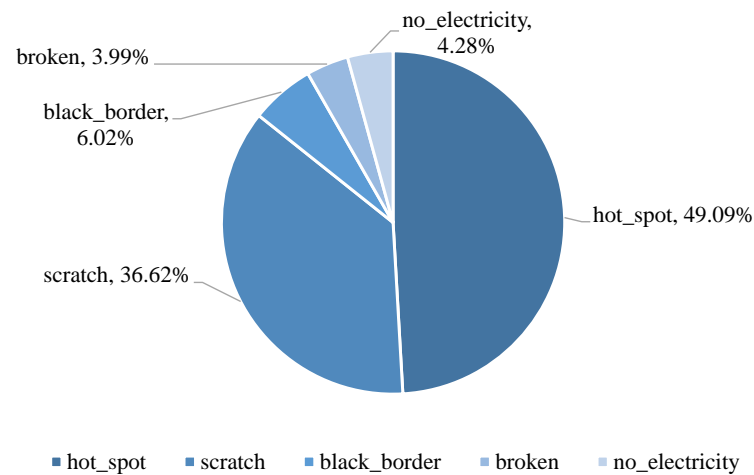
The raw image data pixel size was 5800 \times 3504 pixels, while the average size of the scratches was about 4 \times 32 pixels. The average length of the black edges was about 4 \times 37 pixels, the average length of the broken edges was about 104 \times 210 pixels, the average size of the hot spots was about 152 \times 210 pixels, and the average size of the defective unpowered cell was 356 \times 478 pixels. The size of each type of defect was not uniform, and the size of the defect was less than 0.08% of the whole image. If the original size image were used as the data for training, smaller targets would be detected with lower accuracy or even difficult to detect. Therefore, the original image is preprocessed in this paper, and the image size is changed to 600 \times 600 pixels by compression and cropping operations. Figure 7 compares the performance of each defect before and after data preprocessing (mAP). It can be found that preprocessing improves the training efficiency on the one hand and makes the influence of some tiny target (scratch) noise reduced on the other hand, which effectively improves the accuracy of tiny target detection and increases the generalization ability of the training network.

After preprocessing, we finally obtained 1108 defect images of the PV panel surface. We increased the size of the labeled boxes for each defect type, effectively improving the detection of small and fuzzy targets. The final dataset was sequentially labeled using the LabelImg labeling software concerning the VOC2007 dataset format and converted to the XML format required for training. LabelImg is a labeling tool written in Python for deep learning image dataset production, which was used to label the information of the category name and the location of targets in the images. There were 886 images in the training set and 222 in the validation set.

There were 4235 defective targets in 1108 images of the PV panel surface. Figure 5 shows that hot spots accounted for the highest percentage among the five types of defects, at 49.09%. The tiny target scratches accounted for 36.62%, and the blurred targets black border and broken cells accounted for 6.02% and 3.99%, respectively.

Table 1. Defect sample diagram.

Name of Defect	Description	Image Style			
broken	Photovoltaic panels with broken areas				
hot_spot	Photovoltaic panels have obvious bright spot areas				
black_border	Photovoltaic panels with black or gray border areas				
scratch	Photovoltaic panels with scratched areas				
no_electricity	Photovoltaic panels have non-electricity and show black areas				

**Figure 5.** PV Multi-Defect dataset annotation distribution.

4.2. Baseline Introduction

This work used the confusion matrix for supervised learning as an evaluation metric. The resulting evaluation criteria were the *Precision*, *Recall*, and *mAP* values [31].

The *Recall* is for the original sample and indicates how many positive sample cases were correctly predicted in the model, which is calculated in Equation (6):

$$R = \frac{TP}{TP + FN} \times 100\%. \quad (6)$$

where R denotes the *Recall* rate. *Precision* is for the final prediction result, which indicates how many of the samples with positive predictions were really positive samples, calculated as Equation (7).

$$P = \frac{TP}{TP + FP} \times 100\%. \quad (7)$$

where P denotes the *Precision*. In the actual experiments, since both the *Recall* and *Precision* were maintained at a high level, a parameter was needed to combine *Recall* and *Precision*, i.e., the performance of the algorithmic network is measured in terms of Mean Average Precision (mAP), which applies to multitarget detection and is denoted as Equation (8).

$$mAP = \frac{\sum_{k=1}^N P(k) \Delta R(k)}{C}. \quad (8)$$

In the above equation, N denotes the number of samples in the validation set, $P(k)$ denotes the magnitude of precision P when k targets are detected simultaneously, and $\Delta R(k)$ denotes the change in *recall* when the number of detected samples changes from $k - 1$ to k . C denotes the number of classes of the model.

4.3. Experiment Settings

The experiments were conducted with 48 GB RAM and an RTX3090 graphics card with Pytorch and CUDA versions 1.8 and 11.1, respectively. This study employs Adam as the optimizer according to the pre-training weight of the COOC data set to address the issue of inadequate data and increase learning speed and accuracy. The batch size is 16, and the learning rate is set at 0.001. As demonstrated in Figure 6, where the training and verification losses of the photovoltaic panel defect detector converge after 500 epochs, we monitor the loss of the bounding box by monitoring the loss of the photovoltaic panel defect detector in order to avoid overfitting. The monitoring line demonstrates unequivocally that neither overfitting nor underfitting are present in the learned model. In addition, the validation set was held-out until the last test.

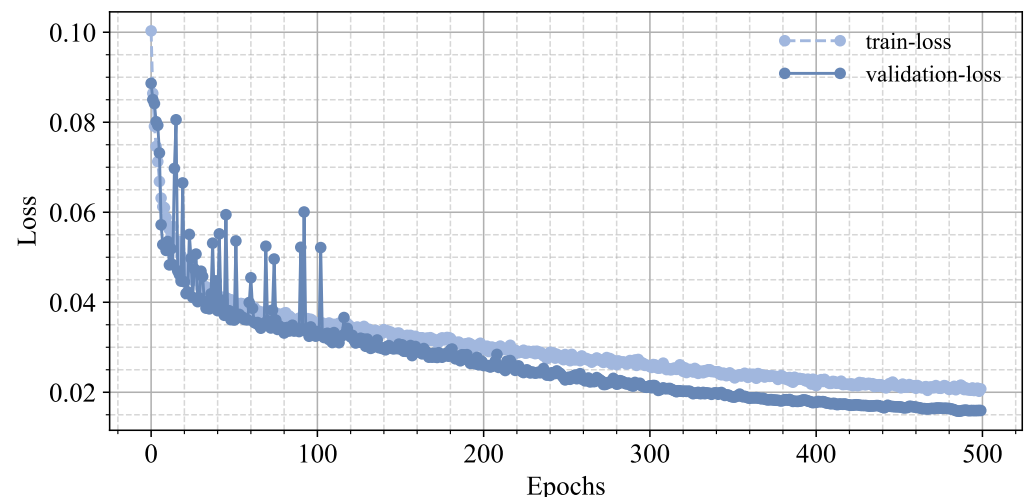


Figure 6. Training and validation losses for PV panel defect detector.

To make the model's performance optimal, the *Recall* and *Precision* values during the training process are monitored in this paper. Table 2 shows their changes during the training of the model, and according to the results, the result of 500 training rounds is chosen as the model for subsequent test experiments in this paper.

4.4. Comparison with Other Methods on the Multi-Defect Dataset

We conducted a set of comparative experiments with the same database and experimental setup, and all methods were retrained. In addition, the validation set was held-out

until the last test. Comparing the methods proposed in this paper with the five techniques discussed in the literature, which are the YOLOv3-based method from Tommaso et al. [15], the Faster-RCNN-based method from Girshick et al. [39], the SVM-based method from Mantel et al. [13], the Mask-RCNN-based method from Almazroue et al. [40], and the SSD-based method from Ren et al. [41], the results showed that our performance on the Multi-Defect dataset was much better than the other models. The specific mAP performance is listed in Table 3.

Table 2. The Performance of GBH-YOLOv5, based on 95% confidence interval.

Epoch	Recall (%)	Precision (%)
0	0	0
50	80.8 ± 0.05	62.9 ± 0.06
100	79.6 ± 0.05	73.1 ± 0.06
150	87.0 ± 0.04	77.9 ± 0.06
200	84.6 ± 0.05	88.1 ± 0.04
250	86.3 ± 0.05	93.0 ± 0.03
300	91.7 ± 0.04	90.1 ± 0.04
350	91.9 ± 0.04	90.7 ± 0.04
400	93.3 ± 0.03	93.5 ± 0.03
450	93.8 ± 0.03	94.5 ± 0.03
500	96.4 ± 0.02	93.3 ± 0.03

Table 3. The mAP performance of different methods on the Multi-Defect dataset, based on 95% confidence interval.

Methods	mAP (%)
Tommaso et al. [15]	57.9 ± 0.07
Girshick et al. [39]	69.3 ± 0.06
Mantel et al. [13]	45.3 ± 0.07
Almazroue et al. [40]	51.2 ± 0.06
Ren et al. [41]	30.8 ± 0.06
Proposed GBH-YOLOv5	97.8 ± 0.02

4.5. Ablation Studies

We analyzed the performance of each component on the PV Multi-Defect dataset, and the impact of each component is presented in Table 4.

Table 4. Ablation Study of PV Multi-Defect datasets, based on 95% confidence interval.

Methods	Description	mAP (%)	Precision (%)	Recall (%)
YOLOv5s	YOLOv5s	78.1 ± 0.06	83.2 ± 0.05	73.4 ± 0.06
YOLOv5-1	YOLOv5s + BottleneckCSP	94.2 ± 0.03	88.2 ± 0.04	90.5 ± 0.04
YOLOv5-2	YOLOv5s + BottleneckCSP + extra prediction	97.1 ± 0.02	93.4 ± 0.03	94.6 ± 0.03
GBH-YOLOv5	YOLOv5s + BottleneckCSP + extra prediction + GhostConv	97.8 ± 0.02	96.4 ± 0.02	93.3 ± 0.02

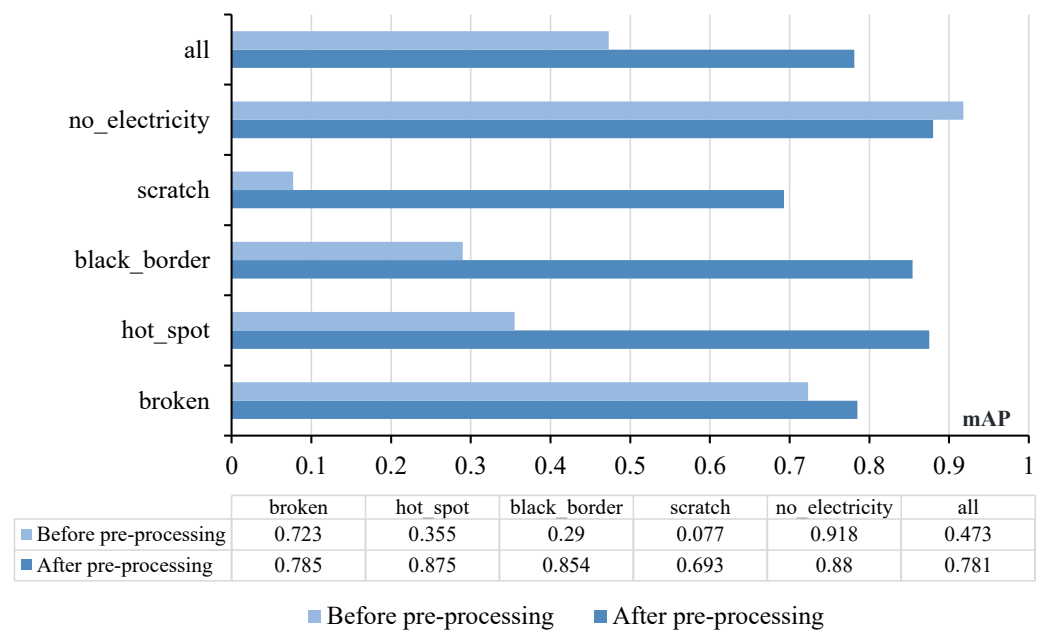


Figure 7. The comparison of each defect's performance in the PV panel before and after data pre-processing.

(1). The effect of BottleneckCSP. After replacing the C3 residual module with the BottleneckCSP module, the mAP values of the model were significantly improved. As shown in Table 5, the performance on all defects was considerably enhanced. The problems of tiny target missed detection and detection accuracy were solved to a large extent. Therefore, the improvement of the BottleneckCSP module was excellent.

Table 5. Performance comparison of the YOLOv5 model for each category on the test set.

Methods	Broken (%)	Hot_Spot (%)	Black_Border (%)	Scratch (%)	No_Electricity (%)
YOLOV5s	78.5 ± 0.05	87.8 ± 0.04	85.4 ± 0.02	69.3 ± 0.06	88.0 ± 0.04
YOLOv5-1	99.5 ± 0.01	97.2 ± 0.02	96.4 ± 0.02	95.6 ± 0.02	97.7 ± 0.02
YOLOv5-2	99.5 ± 0.01	98.4 ± 0.02	96.7 ± 0.02	96.4 ± 0.02	98.9 ± 0.01
GBH-YOLOv5	99.5 ± 0.01	97.5 ± 0.02	97.2 ± 0.02	97.4 ± 0.02	98.0 ± 0.02

(2). The effect of the extra prediction head. As shown in Table 6, the addition of the tiny target prediction head increased the number of network layers in YOLOv5-2 from 224 in YOLOv5s to 290. Even though the computation and the number of parameters were increased, the improvement in the mAP was significant. As shown in Figure 8, GBH-YOLOv5 performed well in detecting tiny targets, so it was worth sacrificing some of the computation.

Table 6. The summary of the different models and the average elapsed time on the test set.

Methods	Model Layers	Number of Parameters (10 ⁴)	Average Time Consuming (s)
YOLOv5s	224	7.06	0.484
YOLOv5-1	228	7.15	0.658
YOLOv5-2	290	7.72	0.695
GBH-YOLOv5	270	7.24	0.587

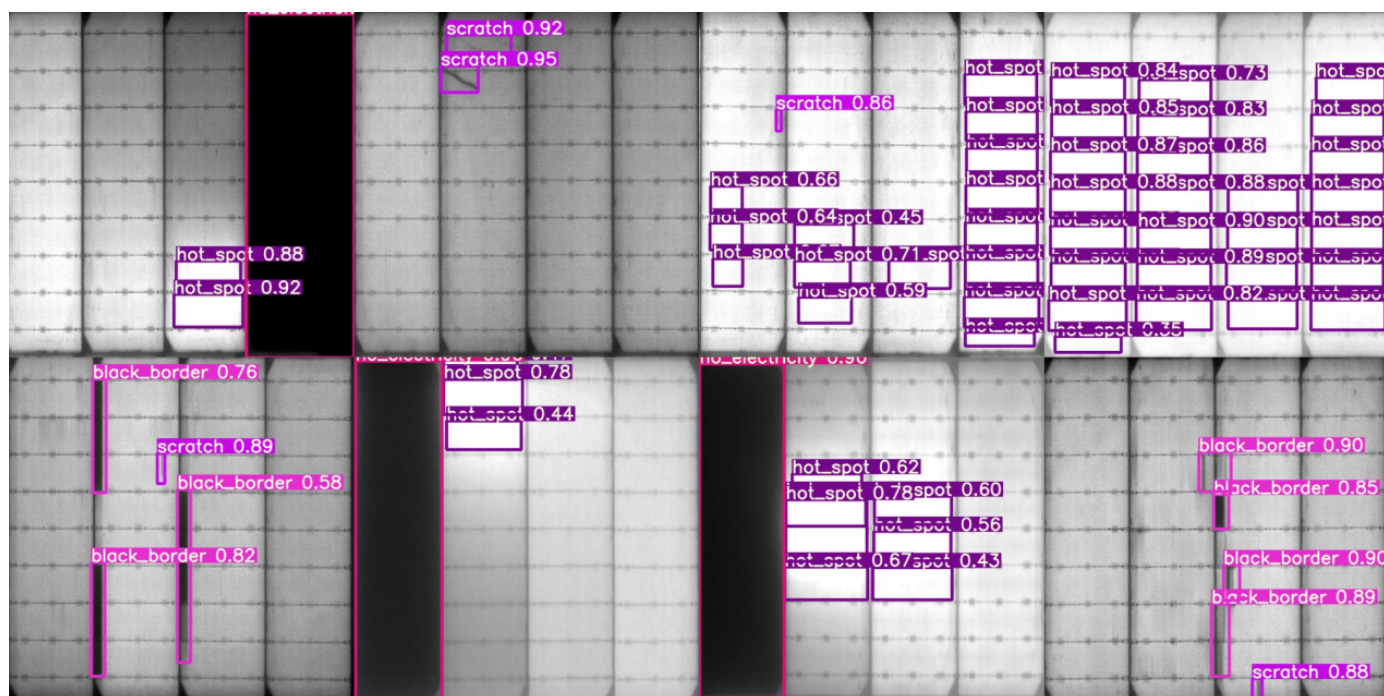


Figure 8. Surface defect identification results of PV panel.

(3). **The effect of GhostCov.** The use of Ghost convolution instead of regular convolution resulted in a reduction in the number of layers in the network of GBH-YOLOv5 from 290 to 270, and the elapsed time on the test set was reduced by 0.108 s per image on average while still maintaining an excellent mAP.

(4). **The effect of the model ensemble.** This paper lists the mAP values of the four models for different categories on the same test set. They are compared with the final integrated model (GBH-YOLOv5) in Table 5, where the GBH-YOLOv5 achieved a relatively balanced result in maintaining accuracy and the time duration.

5. Conclusions

In this paper, we proposed an approach named Ghost convolution with BottleneckCSP and tiny target prediction head incorporating YOLOv5 (GBH-YOLOv5) for PV panel defect detection. To ensure better accuracy on multiscale targets, the BottleneckCSP module was introduced to add a prediction head for tiny target detection to improve the phenomenon of missed detection of tiny defects, and it used Ghost convolution to improve the model inference speed and reduce the number of parameters. First, the original image was compressed and cropped to enlarge the defect size physically. Then, the processed images were input into GBH-YOLOv5, and the depth features were extracted through network processing based on Ghost convolution, the application of the BottleneckCSP module, and the prediction head of tiny targets. Finally, the extracted features were classified by a Feature Pyramid Network (FPN) and a Path Aggregation Network (PAN) structure. Meanwhile, we compared our method with state-of-the-art methods to verify its effectiveness. The proposed PV panel surface-defect detection network improved the mAP performance by at least 27.8%. As the addition of modules makes the number of parameters of the model increase and the volume of the model become larger, the selected dataset is grayscale processed and may generate some errors when detecting PV panel defects in a natural production environment. Possible future research directions include the use of lightweight networks with better real-time performance or the direct use of RGB images for PV panel defect detection.

Author Contributions: Conceptualization, Z.W. and L.L.; methodology, Z.W. and L.L.; software, L.L. and T.Z.; validation, Z.W. and L.L.; formal analysis, Z.W. and L.L.; investigation, Z.W. and L.L.; writing—original draft preparation, L.L. and T.Z.; writing—review and editing, Z.W.; supervision, Z.W.; project administration, Z.W.; funding acquisition, Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: The research work in this paper was supported by the National Natural Science Foundation of China (No. 62177022, 61901165, 61501199), the Collaborative Innovation Center for Informatization and Balanced Development of K-12 Education by MOE and Hubei Province (No. xtzd2021-005), and the Self-determined Research Funds of CCNU from the Colleges' Basic Research and Operation of MOE (No. CCNU22QN013).

Data Availability Statement: Data will be made available on reasonable request.

Acknowledgments: We would like to thank Chunyan Zeng of Hubei University of Technology for her strong support of this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ko, C.C.; Liu, C.Y.; Zhou, J.; Chen, Z.Y. Analysis of subsidy strategy for sustainable development of environmental protection policy. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *349*, 012–018. [\[CrossRef\]](#)
2. Wei, Y.M.; Chen, K.; Kang, J.N.; Chen, W.; Wang, X.Y.; Zhang, X. Policy and Management of Carbon Peaking and Carbon Neutrality: A Literature Review. *Engineering* **2022**, *14*, 52–63. [\[CrossRef\]](#)
3. Almalki, F.A.; Albraikan, A.A.; Soufiene, B.O.; Ali, O. Utilizing Artificial Intelligence and Lotus Effect in an Emerging Intelligent Drone for Persevering Solar Panel Efficiency. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 7741535. [\[CrossRef\]](#)
4. Akram, M.W.; Li, G.; Jin, Y.; Chen, X. Failures of Photovoltaic modules and their Detection: A Review. *Appl. Energy* **2022**, *313*, 118822. [\[CrossRef\]](#)
5. Zeng, C.; Ye, J.; Wang, Z.; Zhao, N.; Wu, M. Cascade Neural Network-Based Joint Sampling and Reconstruction for Image Compressed Sensing. *Signal Image Video Process.* **2022**, *16*, 47–54. [\[CrossRef\]](#)
6. Guerriero, P.; Cuzzo, G.; Daliento, S. Health diagnostics of PV panels by means of single cell analysis of thermographic images. In Proceedings of the 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), Florence, Italy, 7–10 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–6. [\[CrossRef\]](#)
7. Wang, Z.; Wang, Z.; Zeng, C.; Yu, Y.; Wan, X. High-Quality Image Compressed Sensing and Reconstruction with Multi-Scale Dilated Convolutional Neural Network. *Circuits Syst. Signal Process.* **2022**, *39*, 1–24.
8. Rahman, M.M.; Khan, I.; Alameh, K. Potential measurement techniques for photovoltaic module failure diagnosis: A review. *Renew. Sustain. Energy Rev.* **2021**, *151*, 111532. [\[CrossRef\]](#)
9. Mellit, A.; Kalogirou, S. Assessment of machine learning and ensemble methods for fault diagnosis of photovoltaic systems. *Renew. Energy* **2022**, *184*, 1074–1090. [\[CrossRef\]](#)
10. AbdulMawjood, K.; Refaat, S.S.; Morsi, W.G. Detection and prediction of faults in photovoltaic arrays: A review. In Proceedings of the 2018 IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG), Doha, Qatar, 10–12 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–8. [\[CrossRef\]](#)
11. Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B. A Review of Yolo Algorithm Developments. *Procedia Comput. Sci.* **2022**, *199*, 1066–1073. [\[CrossRef\]](#)
12. Lyu, L.; Wang, Z.; Yun, H.; Yang, Z.; Li, Y. Deep Knowledge Tracing Based on Spatial and Temporal Representation Learning for Learning Performance Prediction. *Appl. Sci.* **2022**, *12*, 7188. [\[CrossRef\]](#)
13. Mantel, C.; Villebro, F.; dos Reis Benatto, G.A.; Parikh, H.R.; Wendlandt, S.; Hossain, K.; Poulsen, P.B.; Spataru, S.; Séra, D.; Forchhammer, S. Machine learning prediction of defect types for electroluminescence images of photovoltaic panels. In *Applications of Machine Learning*; Zelinski, M.E., Taha, T.M., Howe, J., Awwal, A.A., Iftekharuddin, K.M., Eds.; SPIE: Bellingham, WA, USA, 2019; Volume, 11139, pp. 1–9. [\[CrossRef\]](#)
14. Wang, Z.; Zuo, C.; Zeng, C. SAE Based Unified Double JPEG Compression Detection System for Web Image Forensics. *Int. J. Web Inf. Syst.* **2021**, *17*, 84–98. [\[CrossRef\]](#)
15. Tommaso, A.D.; Betti, A.; Fontanelli, G.; Michelozzi, B. A multi-stage model based on YOLOv3 for defect detection in PV panels based on IR and visible imaging by unmanned aerial vehicle. *Renew. Energy* **2022**, *193*, 941–962. [\[CrossRef\]](#)
16. Zhu, X.; Lyu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 11–17 October 2021; pp. 2778–2788. [\[CrossRef\]](#)
17. Juan, R.O.S.; Kim, J. Photovoltaic Cell Defect Detection Model based-on Extracted Electroluminescence Images using SVM Classifier. In Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), Fukuoka, Japan, 19–21 February 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 578–582. [\[CrossRef\]](#)

18. Segovia Ramírez, I.; Das, B.; García Márquez, F.P. Fault detection and diagnosis in photovoltaic panels by radiometric sensors embedded in unmanned aerial vehicles. *Prog. Photovolt. Res. Appl.* **2022**, *30*, 240–256. Available online: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/pip.3479> (accessed on 12 November 2022). [\[CrossRef\]](#)
19. Kirubakaran, V.; Preethi, D.M.D.; Arunachalam, U.; Rao, Y.K.S.S.; Gatasheh, M.K.; Hoda, N.; Anbese, E.M. Infrared Thermal Images of Solar PV Panels for Fault Identification Using Image Processing Technique. *Int. J. Photoenergy* **2022**, *2022*, 6427076. [\[CrossRef\]](#)
20. Vergura, S.; Marino, F.; Carpentieri, M. Processing infrared image of PV modules for defects classification. In Proceedings of the 2015 International Conference on Renewable Energy Research and Applications (ICRERA), Palermo, Italy, 22–25 December 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 1337–1341. [\[CrossRef\]](#)
21. Azkona, N.; Llaría, A.; Cúrea, O.; Recart, F. Detection, Characterization and Modeling of Localized Defects and Thermal Breakdown in Photovoltaic Panels from Thermal Images and IV Curves. *Electron. Mater.* **2022**, *3*, 154–172. [\[CrossRef\]](#)
22. Zeng, C.; Yan, K.; Wang, Z.; Yu, Y.; Xia, S.; Zhao, N. Abs-CAM: A Gradient Optimization Interpretable Approach for Explanation of Convolutional Neural Networks. *Signal Image Video Process.* **2022**, *15*, 1–8. [\[CrossRef\]](#)
23. Aouat, S.; Ait-hammi, I.; Hamouchene, I. A new approach for texture segmentation based on the Gray Level Co-occurrence Matrix. *Multimed. Tools Appl.* **2021**, *80*, 24027–24052. [\[CrossRef\]](#)
24. Jumaboev, S.; Jurakuziev, D.; Lee, M. Photovoltaics Plant Fault Detection Using Deep Learning Techniques. *Remote Sens.* **2022**, *14*, 3728. [\[CrossRef\]](#)
25. Lu, F.; Niu, R.; Zhang, Z.; Guo, L.; Chen, J. A Generative Adversarial Network-Based Fault Detection Approach for Photovoltaic Panel. *Appl. Sci.* **2022**, *12*, 1789. [\[CrossRef\]](#)
26. Kurukuru, V.S.B.; Haque, A.; Tripathy, A.K.; Khan, M.A. Machine learning framework for photovoltaic module defect detection with infrared images. *Int. J. Syst. Assur. Eng. Manag.* **2022**, *13*, 1771–1787. [\[CrossRef\]](#)
27. Bu, C.; Liu, T.; Li, R.; Shen, R.; Zhao, B.; Tang, Q. Electrical Pulsed Infrared Thermography and supervised learning for PV cells defects detection. *Sol. Energy Mater. Sol. Cells* **2022**, *237*, 1–7. [\[CrossRef\]](#)
28. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 779–788. [\[CrossRef\]](#)
29. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 6517–6525. [\[CrossRef\]](#)
30. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.
31. Bochkovskiy, A.; Wang, C.; Liao, H.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
32. Su, S.; Yuan, D.; Wang, Y.; Ding, M. Fine Grained Feature Extraction Model of Riot-related Images Based on YOLOv5. *Comput. Syst. Sci. Eng.* **2023**, *45*, 85–97. [\[CrossRef\]](#)
33. Li, C.; Li, L.; Jiang, H.; Weng, K.; Geng, Y.; Li, L.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W.; et al. YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications. *arXiv* **2022**, arXiv:2209.02976.
34. Wang, C.; Bochkovskiy, A.; Liao, H.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv* **2022**, arXiv:2207.02696.
35. Wang, Z.; Yang, Y.; Zeng, C.; Kong, S.; Feng, S.; Zhao, N. Shallow and Deep Feature Fusion for Digital Audio Tampering Detection. *EURASIP J. Adv. Signal Process.* **2022**, *2022*, 69. [\[CrossRef\]](#)
36. Zeng, C.; Zhu, D.; Wang, Z.; Wu, M.; Xiong, W.; Zhao, N. Spatial and Temporal Learning Representation for End-to-End Recording Device Identification. *EURASIP J. Adv. Signal Process.* **2021**, *2021*, 41. [\[CrossRef\]](#)
37. Wang, C.; Liao, H.M.; Wu, Y.; Chen, P.; Hsieh, J.; Yeh, I. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR Workshops 2020, Seattle, WA, USA, 14–19 June 2020; Computer Vision Foundation/IEEE: Piscataway, NJ, USA, 2020; pp. 1571–1580. [\[CrossRef\]](#)
38. Han, K.; Wang, Y.; Tian, Q.; Guo, J.; Xu, C.; Xu, C. GhostNet: More Features From Cheap Operations. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 14–19 June 2020; pp. 1577–1586. [\[CrossRef\]](#)
39. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
40. Almazrouei, H.I.A.A. Enhance PV Panel Detection Using Drone Equipped With RTK. In Proceedings of the ASME 2020 International Mechanical Engineering Congress and Exposition, Virtual, Online, 16–19 November 2020; Volume 7A. [\[CrossRef\]](#)
41. Ren, Y.W.; Yu, Y.; Li, J.H.; Zhang, W.H. Design of photovoltaic hot spot detection system based on deep learning. *J. Phys. Conf. Ser.* **2020**, *1693*, 012075. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.