

Article

Salt Deposit Detection on Offshore Photovoltaic Modules Using an Enhanced YOLOv8 Framework

Gang Li ^{1,2}, Shuqing Wang ^{1,3}, Bo Liu ², Mingqiang Xu ^{1,3,*}, Zhenhai Liu ² and Haoge Wang ¹

¹ The College of Engineering, Ocean University of China, Qingdao 266100, China; gli@cpecc.net (G.L.); shuqing@ouc.edu.cn (S.W.); wanghaoge@stu.ouc.edu.cn (H.W.)

² China Power Engineering Consulting Group Co., Ltd., Beijing 100120, China; bliu@cpecc.net (B.L.); liuzhenhai1201@163.com (Z.L.)

³ State Key Laboratory of Coastal and Offshore Engineering, Ocean University of China, Qingdao 266100, China

* Correspondence: mingqiang.xu@ouc.edu.cn

Abstract

To address the challenges of low detection efficiency and limited accuracy in identifying contamination on offshore photovoltaic platforms, this study proposes an enhanced YOLOv8-based algorithm for detecting salt deposit on photovoltaic modules. The SimAM parameter-free attention mechanism is integrated at the end of the backbone network and within the neck layers to improve feature representation of salt deposits under complex environmental conditions, thereby enhancing detection accuracy. In addition, the WIoU loss function is employed in place of the original CIoU loss to alleviate harmful gradients caused by low-quality data and to strengthen the generalization capability of the model. A dedicated dataset of salt accumulation images from offshore photovoltaic panels is constructed to support this targeted detection task. Experimental results demonstrate that the proposed algorithm achieves an mAP50 of 85.8%, a 3% improvement over YOLOv8, while maintaining a detection speed of 67 frames per second. These findings confirm that the proposed approach meets both the accuracy and efficiency requirements for automated detection of salt deposition on offshore photovoltaic modules.

Keywords: object detection; deep learning; offshore photovoltaics; YOLOv8; attention mechanism

1. Introduction

Compared with land-based photovoltaic power plants, offshore photovoltaic modules take advantage of vast ocean space, thereby overcoming land constraints. The natural cooling effect of seawater further enhances the power generation efficiency of photovoltaic installations. These benefits have gradually positioned offshore photovoltaic modules as a promising pathway for industry expansion and energy transition [1]. However, offshore photovoltaic modules are particularly vulnerable to contamination from sea salt deposition, seabird droppings, and marine biological attachments. Such pollutants reduce energy conversion efficiency, cause power losses, and in severe cases induce hotspots, leading to module damage and compromising long-term system stability [2]. Therefore, timely and accurate detection of surface contamination on offshore photovoltaic modules is of great practical significance for ensuring efficient and reliable system operation.

Current approaches for detecting debris on photovoltaic modules primarily include manual inspection and automated detection methods. Manual inspection suffers from high labor costs, low efficiency, subjectivity, and limited accuracy in large-scale deployments [3].



Academic Editor: Santiago Silvestre

Received: 6 December 2025

Revised: 30 December 2025

Accepted: 3 January 2026

Published: 6 January 2026

Copyright: © 2026 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.

As a result, automated detection technologies, particularly those based on unmanned aerial vehicles (UAVs), have emerged as effective alternatives [4–7]. For example, Tsanakas et al. [8] proposed a hotspot detection method using standard infrared imagery combined with the Canny edge operator, while Ngo et al. [9] applied infrared imaging with K-means clustering for contour extraction and hotspot identification. However, these traditional vision-based methods struggle with robustness in complex backgrounds where key edge information may be obscured. To address such limitations, deep learning has been introduced into photovoltaic defect detection [10–12]. Greco et al. [13] pioneered the use of the YOLO algorithm for aerial debris detection, achieving effective module segmentation and hotspot localization. Zhao et al. proposed a GP-YOLO defect detector that employed reversible column networks and global attention mechanism. GP-YOLO improves mAP@0.5 and mAP@0.5:0.95 by 2.47% and 5.25%, while reducing parameters and GFLOPs by 22.52% and 20.77% compared with the baseline. Li et al. [14] proposed YOLO-LitePV, leveraging multi-scale attention, channel pruning, and knowledge distillation to achieve efficient and accurate PV defect detection. Zhu et al. [15] proposed C2DEM-YOLO, which integrates a deep–shallow C2Dense feature extractor, cross-space multi-scale attention, and an Inner-CIoU regression strategy to enhance EL-based PV defect detection, achieving notable improvements in mAP over YOLOv5 and YOLOv8 on both the PVEL-AD and a self-built dataset. Gan et al. [16] introduced a lightweight transformer-based PV defect detection model that incorporates a PConv-enhanced backbone, a local–global collaborative attention mechanism, and a multipath hybrid enhancement module, achieving notable gains in AP and small-target performance on the PVEL-AD dataset.

Despite these advances, most existing studies have concentrated on land-based photovoltaic systems [3,17,18], where salt deposition is not a major concern. In contrast, offshore photovoltaic modules face severe challenges from salt deposits, which crystallize on the module surface during operation. Salt crystallization occurs when water vapor carrying dissolved salts from seawater condenses on the glass surface, leaving solid deposits. These deposits not only block microscopic pores on the glass, reducing light transmission and photoelectric conversion efficiency, but also accelerate corrosion and shorten the service life of the modules. Salt deposit contamination is particularly critical for coastal and offshore solar installations, where high salinity and humidity intensify degradation. One major consequence is potential-induced degradation (PID), driven by sodium ions present in salt deposits. As deposits accumulate on the glass surface, conductive channels may form, compromising the module's resistance to PID. Moreover, high humidity promotes water-film formation on the module surface, facilitating sodium-ion migration from the glass bulk to the surface. Dust and other atmospheric impurities dissolve into this water film, reducing surface resistivity and further aggravating degradation. The combined effects of salt deposits and humidity therefore pose a serious threat to the performance and reliability of offshore photovoltaic modules. Given the significant impact of salt crystallization, effective monitoring and mitigation strategies are urgently needed. This study addresses this challenge by employing advanced computer vision techniques to detect salt deposition on offshore photovoltaic module surfaces, enabling timely maintenance and ensuring reliable system operation.

YOLOv8, as one of the most advanced object detection frameworks, offers notable advantages in accuracy, speed, and scalability, and has been widely adopted across diverse visual detection tasks. However, when applied to offshore photovoltaic modules, several challenges emerge. First, salt deposits often exhibit irregular shapes, low contrast, and large-scale variations under complex marine conditions, making it difficult for the baseline YOLOv8 to capture fine-grained features effectively. Second, the network may fail to adequately focus on the subtle texture and edge information characteristic of salt contami-

nation, leading to reduced detection precision. Finally, the CIoU loss function employed in YOLOv8 is sensitive to low-quality or noisy samples, which can introduce harmful gradients during training and undermine the robustness and generalization performance of the model. Therefore, although YOLOv8 provides a strong foundation for salt deposit detection, its limitations in handling the distinctive characteristics of offshore photovoltaic modules highlight the need for targeted improvements.

To address these limitations, this study introduces two tailored enhancements. The SimAM parameter-free attention mechanism is incorporated at the end of the backbone and within the neck to strengthen the representation of salt deposit features under challenging environmental conditions, thereby improving detection accuracy. Additionally, the WIoU loss function is adopted in place of the original CIoU loss to reduce the influence of low-quality samples and enhance model generalization. To support training and evaluation, a dedicated salt deposit dataset is constructed by combining real-world field data with simulated samples. These improvements together enable intelligent, accurate, and efficient detection of salt deposition on offshore photovoltaic modules.

2. Methodology

2.1. YOLOv8 Baseline Architecture

YOLO (You Only Look Once) is an object detection algorithm capable of performing object localization and classification simultaneously within a single neural network [19]. Unlike traditional multi-stage detection methods, YOLO conducts detection through a single forward pass, thereby achieving high real-time performance. This characteristic has made it widely applicable in various fields, particularly industrial inspection. As a representative single-stage detection framework, YOLO has undergone multiple iterations. Released in 2023, YOLOv8 inherits and optimizes the advantages of its predecessors, achieving state-of-the-art performance in both accuracy and speed [20]. YOLOv8 provides five model variants: n, s, m, l, and x. As network complexity increases, detection accuracy improves at the cost of slower inference speed. Considering the hardware constraints and real-time requirements of unmanned aerial vehicle (UAV) platforms for offshore photovoltaic detection, this study adopts the YOLOv8 model as the baseline due to its lightweight design and high computational efficiency, and further improves it for the task of offshore photovoltaic pollution detection.

As illustrated in Figure 1, the YOLOv8 architecture consists of four main components: Input, Backbone, Neck, and Head.

Input: Preprocessing operations, including Mosaic data augmentation, adaptive image scaling, and grayscale padding, are applied to enrich training samples and enhance model robustness and generalization.

Backbone: The backbone network is responsible for feature extraction and is composed of Conv and C2f modules, with a Spatial Pyramid Pooling Fast (SPPF) layer at the tail. Through convolution and pooling operations, it captures salient features and improves representation capacity.

Neck: The neck aggregates multi-scale information by combining Feature Pyramid Network (FPN) and Path Aggregation Network (PAN) structures. By fusing shallow and deep features through upsampling and downsampling, it enhances semantic understanding and localization precision.

Head: The detection head employs a decoupled structure with separate branches for classification and localization. This design enables more accurate category prediction and bounding-box regression.

In terms of optimization, YOLOv8 introduces a dynamic sample allocation strategy during training. The loss function combines Binary Cross-Entropy (BCE) loss for classifica-

tion, Distribution Focal Loss (DFL), and Complete IoU (CIoU) loss for regression, ensuring balanced learning and improving overall detection performance.

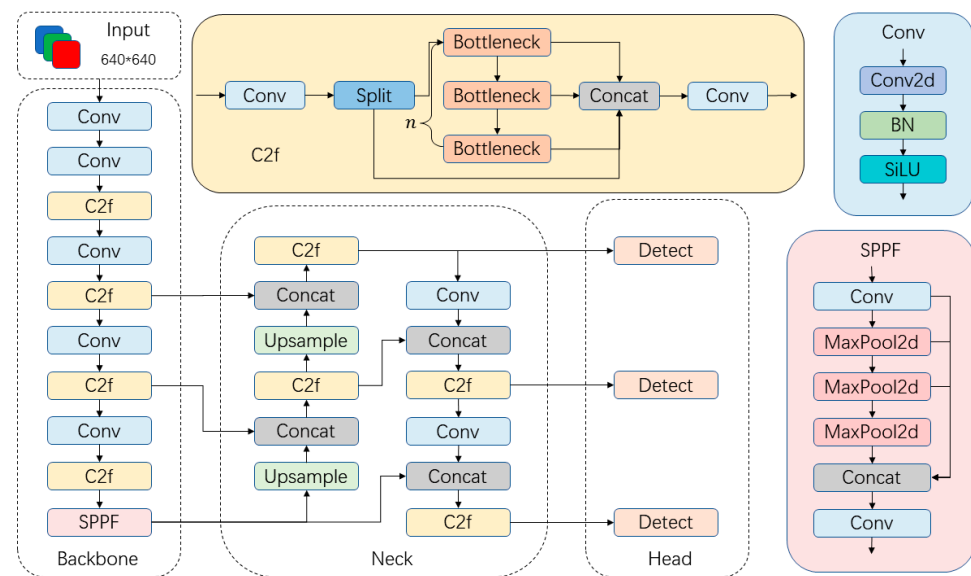


Figure 1. Architecture of the YOLOv8 detection network.

2.2. Proposed Improved YOLOv8 Method

2.2.1. Structural Improvements to the YOLOv8 Network

Building upon the earlier YOLO series, YOLOv8 has been expanded and optimized through the introduction of new features and improvements, enhancing the algorithm's accuracy and flexibility. However, when conducting offshore photovoltaic debris detection in complex marine environments, the following issues may arise: surface reflections causing noise interference, drone aerial images being blurry or distorted, and numerous small or irregular debris making objects difficult to identify. These issues can weaken the model's recognition and detection capabilities, potentially leading to false positives or false negatives. To improve the accuracy of offshore photovoltaic debris detection, this study introduces structural modifications at both the backbone and neck to strengthen feature extraction and multi-scale representation. The overall design philosophy is to enhance discriminative capability without increasing model size, maintaining compatibility with UAV-based real-time inspection.

First, a parameter-free SimAM attention module is integrated into the backbone. The SPPF output is a critical stage where high-level semantic features converge, yet these features may be degraded in offshore scenes due to image blur or local noise. By inserting SimAM after the SPPF layer, the model is guided to emphasize informative neurons while suppressing irrelevant responses, thereby improving the robustness of global feature representations.

Second, the neck structure is further enhanced by embedding SimAM between the feature fusion pathways of the FPN–PAN architecture. Offshore debris often appears as small, scattered particles distributed across multiple scales; thus, multi-level contextual refinement is essential for accurate detection. The attention-enhanced fusion promotes consistency and selectivity across scales, enabling the model to better capture subtle contamination patterns on the photovoltaic surface.

Third, at the detection head level, the original CIoU regression loss is replaced with WIoU to improve localization stability under complex offshore imaging conditions. While this modification is detailed in Section 2.2.3, it forms an integral part of the structural improvement strategy because it alters how the detection head optimizes bounding-box

geometry throughout training. Together with SimAM-enhanced features, WIoU contributes to a more reliable overall detection pipeline.

Through the coordinated integration of these enhancements, the improved YOLOv8 architecture strengthens both feature abstraction and geometric regression without increasing computational cost. The complete structure of the modified network is shown in Figure 2, and the subsequent subsections provide detailed descriptions of the incorporated modules.

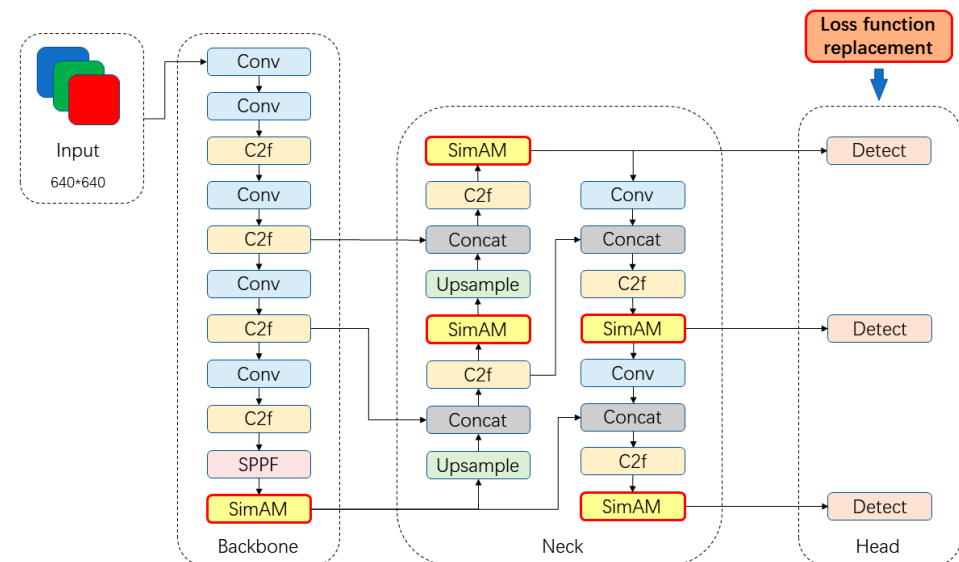


Figure 2. Architecture of the improved YOLOv8 network.

2.2.2. SimAM Attention Mechanism

Attention mechanisms play an increasingly important role in convolutional neural networks (CNNs), enabling models to focus on informative regions and suppress irrelevant background. However, most existing attention modules, such as channel-only or spatial-only mechanisms, require additional parameters and increase computational complexity. To address this limitation, the SimAM (simple parameter-free attention module) [21] is adopted in this study. SimAM is a lightweight, parameter-free attention mechanism inspired by neuronal linear separability in neuroscience, and it assigns importance weights to neurons based on an energy function.

Given an input feature map $\mathbf{X} \in \mathbb{R}^{C \times H \times W}$, SimAM operates at the neuron level. For a target neuron value t within a given channel, SimAM formulates a binary classification problem that separates the target neuron from the remaining $M - 1$ neurons in the same channel. The corresponding energy function is defined as:

$$e_t(w_t, b_t) = (1 - (w_t t + b_t))^2 + \frac{1}{M-1} \sum_{i \neq t} (-1 - (w_t x_i + b_t))^2 + \lambda w_t^2, \quad (1)$$

where w_t and b_t denote the weight and bias of a linear classifier, x_i represents the remaining neurons in the same channel, and $M = H \times W$ is the total number of neurons per channel. The factor $\frac{1}{M-1}$ is applied only to the summation term to normalize the contribution of non-target neurons, while the term λw_t^2 acts as an L_2 regularization to ensure numerical stability of the closed-form solution.

To obtain a closed-form solution, SimAM assumes that the non-target neurons x_i within the same channel follow a common distribution. Accordingly, their statistics are approximated using the global mean $\hat{\mu}$ and variance $\hat{\sigma}^2$, computed over all neurons in the

same channel excluding the target neuron t . Under this assumption, minimizing the energy function with respect to w_t and b_t yields the minimum energy value:

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda}, \quad (2)$$

The importance of neuron t is inversely proportional to its minimum energy. Therefore, the SimAM attention weight is explicitly computed as:

$$W_t = \text{sigmoid}\left(\frac{1}{e_t^*}\right), \quad (3)$$

By applying this operation to all neurons, SimAM generates a three-dimensional attention map $\mathbf{W} \in \mathbb{R}^{C \times H \times W}$. The final reweighted feature map is obtained via element-wise multiplication:

$$\tilde{\mathbf{X}} = \mathbf{W} \odot \mathbf{X}, \quad (4)$$

The structure of the SimAM attention module employed in this study is illustrated in Figure 3. Unlike conventional channel-wise or spatial-wise attention mechanisms, SimAM directly outputs a full 3D importance mask without introducing any additional learnable parameters, enabling fine-grained feature enhancement under noisy and low-contrast maritime conditions.

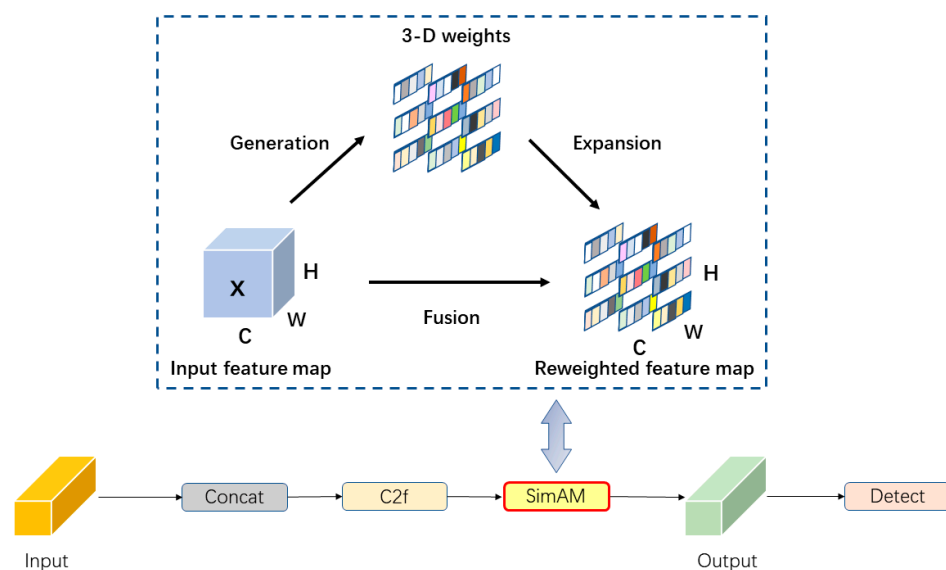


Figure 3. Structure of the SimAM attention module and its integration into the YOLOv8 network.

In this study, SimAM is integrated into the YOLOv8 architecture by placing it after the SPPF module in the backbone to enhance high-level semantic features under challenging conditions. It is further inserted between the feature fusion layers in the neck to refine multi-scale representations and strengthen the detection of small and scattered debris. This integration strategy allows the improved model to emphasize meaningful texture and edge information while suppressing irrelevant noise, thereby achieving more reliable detection performance without incurring additional parameters or computational overhead.

2.2.3. Optimization of the Loss Function

The loss function quantifies the discrepancy between predicted and ground-truth bounding boxes, and its appropriate design is critical for stable and accurate object detection.

YOLOv8 employs the CIoU loss [22] as its regression objective, which accounts for overlap, center-point distance, and aspect-ratio consistency. The CIoU loss is defined as:

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{(x_p - x_{gt})^2 + (y_p - y_{gt})^2}{W_g^2 + H_g^2} + \alpha v, \quad (5)$$

where $\text{IoU} = \frac{B_{gt} \cap B_{pred}}{B_{gt} \cup B_{pred}}$ denotes the intersection-over-union between the predicted box B_{pred} and the ground-truth box B_{gt} , with higher IoU indicating closer alignment. (x_{gt}, y_{gt}) and (x_p, y_p) represent the centers of B_{gt} and B_{pred} , respectively, while W_g and H_g correspond to the width and height of the minimum enclosing rectangle covering both boxes. The term $v = \frac{4}{\pi^2} \left[\arctan\left(\frac{W_{gt}}{H_{gt}}\right) - \arctan\left(\frac{W_p}{H_p}\right) \right]^2$ measures the difference in aspect ratio between the predicted and ground-truth boxes, while $\alpha = \frac{v}{(1 - \text{IoU}) + v}$ determines how strongly this aspect-ratio penalty influences the CIoU loss depending on the IoU level. Here (W_{gt}, H_{gt}) and (W_p, H_p) denote the width and height of B_{gt} and B_{pred} , respectively. The geometric interpretation of these variables is illustrated in Figure 4.

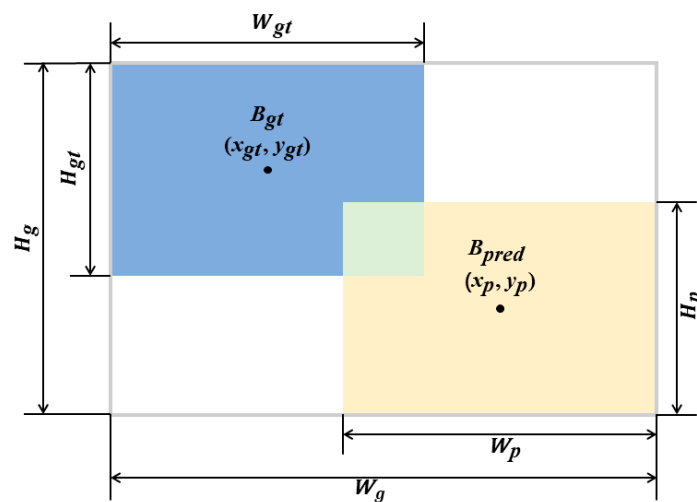


Figure 4. Geometric illustration of variables in CIoU and WIoU losses.

Although CIoU improves bounding-box regression by incorporating geometric constraints, it remains sensitive to localization deviations, particularly in low-quality UAV imagery or when detecting small, irregular offshore debris, often leading to unstable gradients and reduced robustness. To overcome these limitations, this study employs the WIoU (Wise IoU) loss [23] as a replacement for CIoU. WIoU introduces the outlier degree β to quantify the relative quality of a prediction with respect to the running statistics of the IoU loss, which is defined as:

$$\beta = \frac{L_{\text{IoU}}}{\overline{L_{\text{IoU}}}} \in [0, +\infty), \quad (6)$$

where L_{IoU} denotes the IoU-based regression loss of the current predicted-ground-truth pair, and $\overline{L_{\text{IoU}}}$ is the exponential running mean of L_{IoU} updated during training. Based on β , a dynamic non-monotonic focusing mechanism assigns smaller gradient gains to low-quality predictions, effectively suppressing harmful gradients from outliers. The complete WIoU formulation is then expressed as:

$$L_{\text{WIoU}} = L_{\text{IoU}} \cdot \exp \left[\frac{(x_p - x_{gt})^2 + (y_p - y_{gt})^2}{W_g^2 + H_g^2} \right] \cdot \frac{\beta}{\delta \gamma^{(\beta - \delta)}}, \quad (7)$$

where γ and δ are hyperparameters controlling the mapping from the outlier degree β to the gradient gain. In this study, γ and δ are set to 1.9 and 3, respectively, following [23]. The geometric terms (x_p, y_p) , (x_{gt}, y_{gt}) , W_g , and H_g follow the definitions in Equation (5) (CIoU) above.

To further illustrate the behavioral differences between CIoU and WIoU, two representative prediction cases are shown in Figure 5. When the misalignment between the predicted and ground-truth boxes is small, both losses behave similarly; however, under larger deviations, CIoU increases sharply while WIoU grows more moderately, demonstrating its ability to down-weight the influence of low-quality predictions.

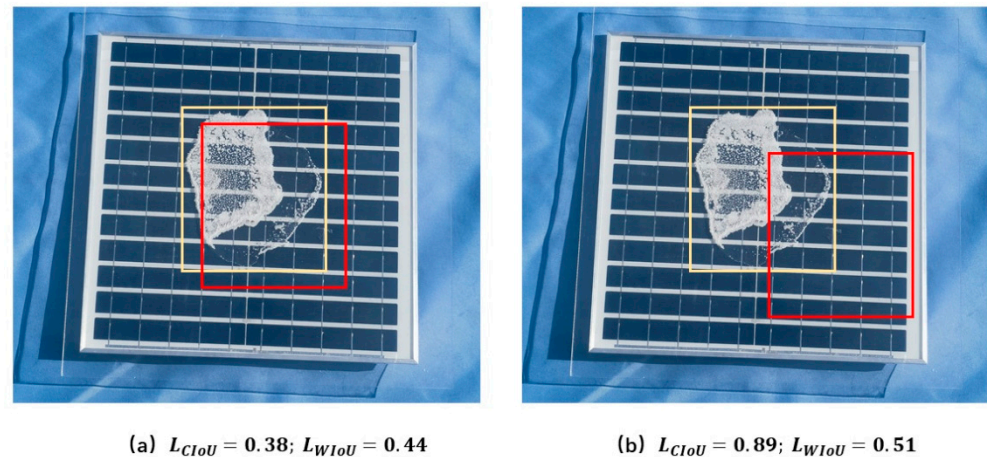


Figure 5. Comparison of CIoU and WIoU losses. Yellow box: ground truth; red box: prediction. (a) Slight misalignment. (b) Large misalignment.

In this study, WIoU is integrated into the regression branch of the YOLOv8 detection head to better address challenges in offshore photovoltaic inspection, including image blur, noise interference, and the prevalence of small debris. This integration enhances the stability of bounding-box optimization, mitigates the impact of low-quality samples, and improves the overall detection accuracy of the proposed model.

3. Salt Deposit Detection Experiments

3.1. Experimental Dataset

Salt deposition is one of the most prevalent contaminants in offshore photovoltaic (PV) operations. It not only reduces the power-generation efficiency by blocking incident irradiance, but also accelerates surface degradation through long-term corrosion. Accordingly, this study targets salt deposits as the primary contamination type for offshore PV inspection.

Due to the lack of publicly available datasets for offshore PV salt deposits, we constructed a dedicated dataset by combining field collection with laboratory-based simulation, as illustrated in Figure 6. First, representative salt-deposit patterns observed in real offshore PV installations were used as references to guide the morphology of simulated samples. Second, a sodium chloride solution with a salinity of 3.5% (approximately the average salinity of seawater) was prepared and randomly applied to PV module surfaces. After evaporation and crystallization, simulated salt deposits were formed. Finally, an experimental imaging platform was established to capture and annotate salt-deposit images for model development.

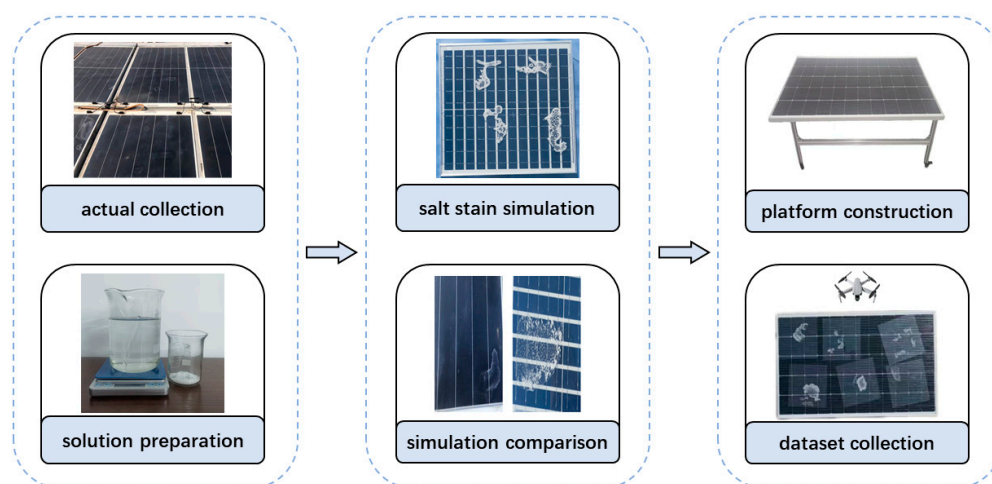


Figure 6. Workflow of the salt deposit simulation experiment.

To mitigate overfitting under limited data and to enhance robustness to environmental variability, multiple simulation trials were conducted under different weather conditions (e.g., sunny, cloudy, and foggy). Data augmentation operations, including cropping, rotation, and mirroring, were applied to increase sample diversity. Consequently, the final dataset contains 500 images in total, comprising 30 images collected from real offshore PV power stations, 84 original laboratory images, and 386 augmented samples generated from the laboratory subset. Among the laboratory images, 36 samples were acquired under a UAV-based imaging setup to reflect drone inspection scenarios, and the remaining samples were obtained from small-scale panel experiments. Representative examples are shown in Figure 7.

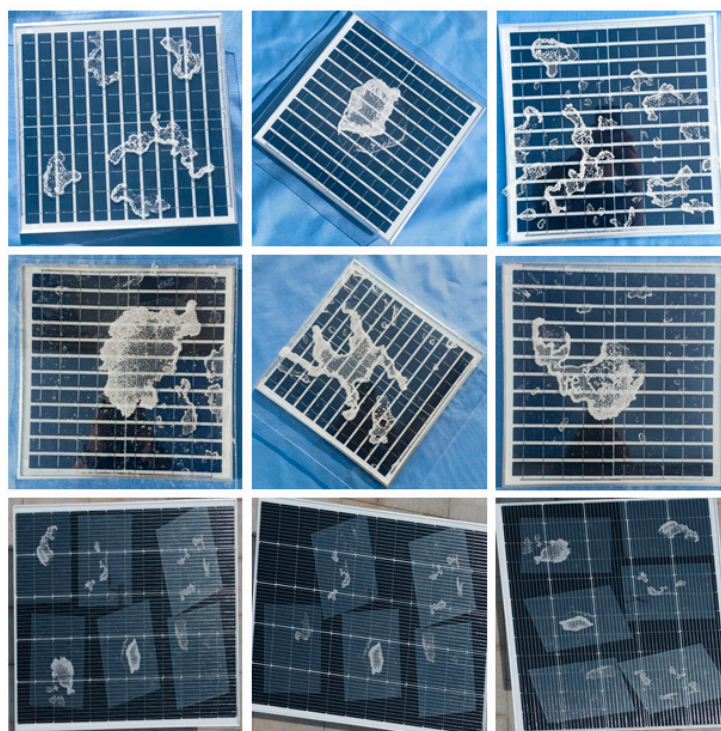


Figure 7. Representative images of offshore photovoltaic panels with salt deposits in the dataset.

After image acquisition, the dataset was randomly split into training, validation, and test sets with a ratio of 8:1:1. All images were annotated using LabelImg 1.8.6 in the YOLO format, where the bounding-box location and category label of salt deposits were

recorded, as illustrated in Figure 8. The annotated labels and corresponding images were then organized into the respective subsets to complete the offshore PV salt-deposit dataset construction. All performance metrics reported in the following sections are evaluated on the held-out test set, which is strictly separated from the training and validation data.

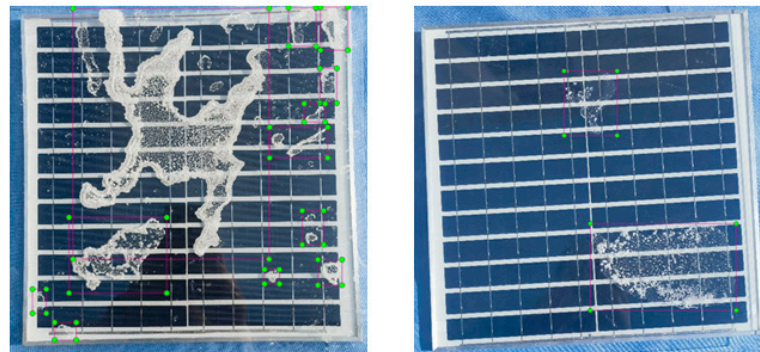


Figure 8. Annotation of salt deposits on offshore photovoltaic panels using Labelling.

3.2. Experimental Setup and Evaluation Metrics

This study was conducted on a 64-bit Windows 11 server with Python 3.8 as the programming environment. The neural network models were implemented using the PyTorch 2.1 deep learning framework and trained with CUDA 12.2 for GPU acceleration. The experimental hardware comprised an Intel® Core™ i7-13700F CPU, an NVIDIA GeForce RTX 4080 GPU, and 16 GB of RAM. For model training, stochastic gradient descent (SGD) was employed as the optimizer. Input images were resized to 640×640 pixels, and training was performed over 300 epochs with a batch size of 32. The initial learning rate was set to 0.01 to facilitate rapid convergence during the early training stages, while subsequent adjustments to the learning rate schedule further optimized training performance.

Model performance was evaluated using four metrics: billions of floating-point operations per second (GFLOPs), model parameter count (Params), frames per second (FPS), and mean average precision (mAP). GFLOPs quantifies the computational complexity of the model, and Params represents the total number of trainable parameters, reflecting model size. Smaller values of GFLOPs and Params indicate lower computational requirements and reduced hardware demands, making the model more suitable for deployment on resource-constrained devices. FPS measures the number of image frames processed per second, which depends on both the algorithm complexity and the underlying hardware.

The mean average precision at a 50% IoU threshold (mAP50) was used to evaluate the detection accuracy and recall performance of the models. Precision (P) and recall (R) are defined as:

$$P = \frac{TP}{TP + FP}, \quad (8)$$

$$R = \frac{TP}{TP + FN}, \quad (9)$$

where TP (true positive) indicates correctly predicted bounding boxes, FP (false positive) indicates predicted boxes without corresponding ground-truth boxes, and FN (false negative) indicates ground-truth boxes missed by the prediction. The mAP50 is then computed by integrating the area under the Precision–Recall curve for each category and averaging across all n categories:

$$mAP = \frac{\sum_{i=1}^n \int_0^1 P(R) dR}{n}, \quad (10)$$

Figure 9 shows the mAP50 curve during training for both the baseline YOLOv8 and the improved YOLOv8 models. Both models exhibit rapid accuracy growth in the

early training stage and converge around 50 epochs, indicating stabilization of model performance. Notably, the improved YOLOv8 model consistently achieves higher mAP50 than the baseline across most iterations, particularly in later stages, demonstrating superior detection accuracy and stability in object detection tasks.

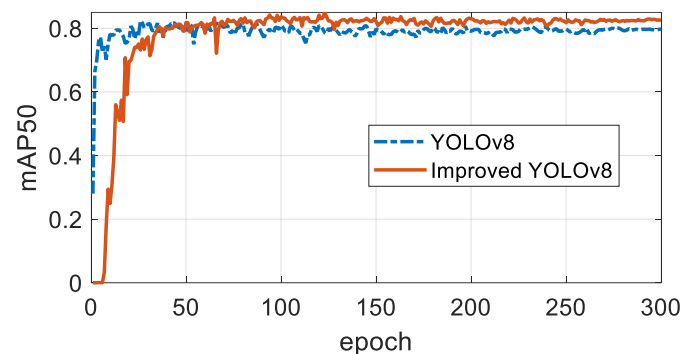


Figure 9. Comparison of mAP50 between the baseline YOLOv8 and the improved YOLOv8 during model training.

3.3. Experimental Results

To enhance the performance of object detection networks, this study systematically improves and evaluates existing attention mechanisms and loss functions. First, the SimAM attention module is introduced and compared with other commonly used attention mechanisms, including the CA and SE modules, to investigate the impact of different attention mechanisms on detection accuracy, model size, and inference speed. Subsequently, the loss function is optimized by adopting the WIoU loss, which is experimentally compared with CIoU, DIoU (Distance-IoU), and GIoU (Generalized IoU) to verify its effectiveness in improving both detection accuracy and speed of the YOLOv8 model. These improvements highlight the critical role of optimizing attention mechanisms and loss functions in enhancing overall object detection performance.

A comparative experiment on attention mechanisms was conducted, with the SimAM module evaluated against the widely used CA (Coordinate Attention) [24] and SE (Squeeze-and-Excitation) [25] modules, both of which have demonstrated strong performance in enhancing feature representation in convolutional neural networks. The results are summarized in Table 1. It can be observed that integrating SimAM increases the mAP50 by 2%, reduces the number of parameters by 0.04 M, decreases GFLOPs by 0.4, and increases FPS by 1.2, demonstrating superior performance in both detection accuracy and speed.

Table 1. Performance comparison of YOLOv8 with different attention mechanisms.

Model	mAP50 (%)	Params (M)	GFLOPs	FPS
YOLOv8	82.8	11.16	28.8	65.3
+CA	83.4	11.33	29.2	60.5
+SE	83.4	11.21	28.7	62
+SimAM	84.8	11.12	28.4	66.5

Next, a comparative study of loss functions was performed, with the WIoU loss compared to CIoU, DIoU (Distance-IoU) [26] and GIoU (Generalized IoU) [27]. The results, presented in Table 2, show that the YOLOv8 model using WIoU achieves an mAP50 of 83.4% and an FPS of 66.2, outperforming other loss functions and confirming the effectiveness of the proposed optimization.

Table 2. Performance comparison of YOLOv8 with different loss functions.

Model	mAP50 (%)	FPS
YOLOv8+CIoU	82.8	65.3
YOLOv8+DIoU	82.4	59.9
YOLOv8+GIoU	83.3	63.5
YOLOv8+WIoU	83.4	66.2

To further validate the improved algorithm, it was compared with the YOLOv8 baseline and representative YOLO networks in related research, including YOLOv5 and YOLOv10, under identical experimental conditions and using the same dataset. The results are shown in Table 3. The improved YOLOv8 model achieves mAP50 values 3.3%, 3%, and 4.7% higher than YOLOv5, YOLOv8 baseline, and YOLOv10, respectively, indicating the highest detection accuracy among the compared models. In terms of model size and inference speed, the improved algorithm contains 11.12 M parameters and achieves 67.3 FPS. Although slightly slower than YOLOv10 when the non-maximum suppression (NMS) post-processing step is removed, the improved model still satisfies real-time detection requirements. Compared with the YOLOv8 baseline, the improved model shows consistent advantages across all metrics. Overall, the proposed improvements achieve a favorable balance between speed and accuracy, demonstrating excellent detection performance for offshore photovoltaic debris and validating the superiority of the improved algorithm.

Table 3. Performance comparison of different YOLO algorithms.

Model	mAP50 (%)	Params (M)	GFLOPs	FPS
YOLOv5	82.5	14.40	15.8	75.0
YOLOv8	82.8	11.16	28.8	65.3
YOLOv10	81.1	8.04	16.6	78.0
Proposed	85.8	11.12	28.4	67.3

3.4. Ablation Experiment

To further verify the effectiveness of the improved YOLOv8 algorithm proposed in this study for salt deposition detection on offshore photovoltaic platforms, ablation experiments were conducted using YOLOv8 as the baseline model. The results are presented in Table 4, where Model A denotes the baseline model with the WIoU loss function replacing the original regression loss, Model B denotes the baseline model enhanced with the SimAM attention mechanism, and Model C incorporates both improvements simultaneously.

Table 4. Ablation study results of the proposed YOLOv8 improvements.

Model	Loss	Attention	P (%)	R (%)	mAP50 (%)	Params (M)	GFLOPs	FPS
YOLOv8			76.1	80.7	82.8	11.16	28.8	65.3
A	✓		79.4	79.2	83.4	11.16	28.8	66.2
B		✓	77.6	80.8	84.8	11.12	28.4	66.5
C	✓	✓	81.7	77.6	85.8	11.12	28.4	67.3

Overall, the ablation results in Table 4 clearly demonstrate the effectiveness of the proposed improvements. Model A shows that replacing the original regression loss with WIoU leads to a consistent improvement in detection accuracy while maintaining model complexity, indicating the benefit of the optimized bounding-box regression strategy. Model B further enhances performance by introducing the SimAM attention mechanism, which strengthens feature representation and improves detection capability without increasing computational cost. By integrating both SimAM and WIoU, Model C achieves the best

overall performance among all variants. Compared with the baseline YOLOv8, Model C improves the mAP50 by 3.0%, reduces the number of parameters by approximately 40 K, decreases GFLOPs by 0.4, and increases inference speed by 2 FPS. These results indicate that the proposed method achieves a favorable balance between detection accuracy and efficiency, while fully satisfying real-time requirements, making it well suited for practical offshore photovoltaic salt-deposit detection.

3.5. Field Test Verification

The photovoltaic project selected for this study is a demonstration of the floating photovoltaic system in Laizhou Bay, Shandong Province, China, where a 0.46 MW array of photovoltaic panels was initially installed offshore. Being a floating system, the panels were exposed to seawater splashes. Later, the panels were removed and placed onshore. As the seawater evaporated, irregularly shaped salt deposits of varying sizes remained on the panel surfaces. To capture these conditions, a total of 30 photographs of the panels were taken using a mobile phone, with random shooting angles and relatively long distances, which made the salt deposits difficult to resolve clearly. Despite the low resolution, manual inspection identified approximately 146 discernible salt deposits across the images, providing a reference for evaluating detection performance.

To validate the effectiveness and accuracy of the improved YOLOv8 algorithm proposed in this study for detecting debris on actual offshore photovoltaic platforms, these 30 images were analyzed. The trained model was applied to detect salt deposits, and representative detection results are presented in Figure 10. The results demonstrate that the improved algorithm can accurately identify numerous small salt deposition objects in real offshore conditions, with no obvious false positives or false negatives. Quantitatively, the model successfully detected 134 out of 146 manually identified salt deposits, achieving an accuracy of 84.9%. These results indicate that the improved YOLOv8 algorithm maintains high detection performance under real-world conditions. The measurements and statistical analysis further suggest that the model exhibits strong robustness and generalization capability, effectively meeting the practical detection requirements for offshore photovoltaic platforms.



Figure 10. Detection results on actual offshore photovoltaic panels.

4. Conclusions

In this study, an improved YOLOv8-based detection framework was developed to address the challenge of accurately identifying salt deposition on offshore photovoltaic panels, where small target scales and complex marine environments often hinder detection performance. The proposed method enhances both feature representation and bounding-box regression through targeted architectural and algorithmic optimizations. To support model development and evaluation, a dedicated offshore photovoltaic salt-deposition dataset was constructed by integrating real-scene UAV imagery with experimentally simulated salt stains generated under controlled laboratory conditions. In addition, real offshore

operational images were used to further verify the practical performance of the proposed algorithm. The main conclusions are as follows:

- (1) Embedding the SimAM parameter-free attention mechanism into both the backbone and neck enhances the ability of the network to extract discriminative features under complex offshore environmental conditions. Experimental results show that SimAM improves the mAP50 of the model while simultaneously reducing the number of parameters and GFLOPs, enabling better accuracy–efficiency balance.
- (2) Substituting the standard bounding-box regression loss with WIoU effectively suppresses the harmful gradients caused by low-quality samples. This improvement enhances the stability of the training process and increases the overall detection accuracy, demonstrating stronger robustness and generalization capability.
- (3) With the combination of both improvements, the proposed model achieves an mAP50 of 85.8%, representing a 3% gain over the YOLOv8 baseline. The improved algorithm also reduces model complexity and increases detection speed, and the final detection performance has been validated through real-measurement tests on the constructed offshore photovoltaic salt deposition dataset.

Future work will focus on deploying the improved model on UAV platforms for real-time inspection of offshore photovoltaic farms, enabling large-scale, automated monitoring in practical engineering applications.

Author Contributions: Conceptualization, S.W. and M.X.; methodology, G.L., S.W., M.X. and H.W.; software, G.L., M.X. and H.W.; validation, S.W., B.L. and Z.L.; formal analysis, G.L., M.X. and H.W.; investigation, B.L. and Z.L.; resources, G.L., B.L. and Z.L.; data curation, H.W.; writing—original draft preparation, G.L. and H.W.; writing—review and editing, S.W. and M.X.; visualization, B.L. and Z.L.; supervision, S.W. and M.X.; project administration, B.L. and Z.L.; funding acquisition, S.W. and M.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Shandong Provincial Natural Science Fund for Excellent Young Scientists Fund Program (Overseas) (2025HWYQ-030), Taishan Young Scholar Expert Program of Shandong Province (TSQN202408083), National Science Foundation of China (52401349), Postdoctoral Innovation Program of Shandong Province (SDCX-ZG-202503049).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Author Gang Li, Bo Liu and Zhenhai Liu were employed by the China Power Engineering Consulting Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Panjwani, B. Assessment of Breakwater as a Protection System against Aerodynamic Loads Acting on the Floating PV System. *Energies* **2024**, *17*, 4873. [\[CrossRef\]](#)
2. Jun, W.; PL, D. Review of Recent Offshore Photovoltaics Development. *Energies* **2022**, *15*, 7462. [\[CrossRef\]](#)
3. Bajc, T.; Kostadinović, D. Potential of usage of the floating photovoltaic systems on natural and artificial lakes in the Republic of Serbia. *J. Clean. Prod.* **2023**, *422*, 138598. [\[CrossRef\]](#)
4. Alsafasfeh, M.; Abdel-Qader, I.; Bazuin, B.; Alsafasfeh, Q.; Su, W. Unsupervised Fault Detection and Analysis for Large Photovoltaic Systems Using Drones and Machine Vision. *Energies* **2018**, *11*, 2252. [\[CrossRef\]](#)
5. Li, X.; Yang, Q.; Chen, Z.; Luo, X.; Yan, W. Visible defects detection based on UAV-based inspection in large-scale photovoltaic systems. *IET Renew. Power Gener.* **2017**, *11*, 1234–1244. [\[CrossRef\]](#)
6. Kuo, C.-F.J.; Chen, S.-H.; Huang, C.-Y. Automatic detection, classification and localization of defects in large photovoltaic plants using unmanned aerial vehicles (UAV) based infrared (IR) and RGB imaging. *Energy Convers. Manag.* **2023**, *276*, 116495. [\[CrossRef\]](#)
7. Pinho, L.S.; Sousa, T.D.; Pereira, C.D.; Pinto, A.M. Anomaly Detection for PV Modules using Multi-modal Data Fusion in Aerial Inspections. *IEEE Access* **2025**, *13*, 88762–88779. [\[CrossRef\]](#)

8. Tsanakas, J.A.; Chrysostomou, D.; Botsaris, P.N.; Gasteratos, A. Fault diagnosis of photovoltaic modules through image processing and Canny edge detection on field thermographic measurements. *Int. J. Sustain. Energy* **2015**, *34*, 351–372. [\[CrossRef\]](#)
9. Ngo, G.C.; Macabebe, E.Q.B. Image segmentation using K-means color quantization and density-based spatial clustering of applications with noise (DBSCAN) for hotspot detection in photovoltaic modules. In Proceedings of the 2016 IEEE Region 10 Conference (TENCON), Singapore, 22–25 November 2016; IEEE: New York, NY, USA, 2016; pp. 1614–1618.
10. Hussain, M.; Khanam, R. In-depth review of yolov1 to yolov10 variants for enhanced photovoltaic defect detection. *Solar* **2024**, *4*, 351–386. [\[CrossRef\]](#)
11. Masita, K.; Hasan, A.; Shongwe, T.; Hilal, H.A. Deep learning in defects detection of PV modules: A review. *Sol. Energy Adv.* **2025**, *5*, 100090. [\[CrossRef\]](#)
12. Hijjawi, U.; Lakshminarayana, S.; Xu, T.; Fierro, G.P.M.; Rahman, M. A review of automated solar photovoltaic defect detection systems: Approaches, challenges, and future orientations. *Sol. Energy* **2023**, *266*, 112186. [\[CrossRef\]](#)
13. Greco, A.; Pironti, C.; Saggese, A.; Vento, M.; Vigilante, V. A deep learning based approach for detecting panels in photovoltaic plants. In Proceedings of the 3rd International Conference on Applications of Intelligent Systems, Las Palmas de Gran Canaria, Spain, 7–9 January 2020; pp. 1–7.
14. Li, X.; Hou, H.; Yu, Y.; Fan, Z.; Zhao, X. YOLO-LitePV: A lightweight detection algorithm for photovoltaic panel defects. *Nondestruct. Test. Eval.* **2025**, 1–28. [\[CrossRef\]](#)
15. Zhu, J.; Zhou, D.; Lu, R.; Liu, X.; Wan, D. C2DEM-YOLO: Improved YOLOv8 for defect detection of photovoltaic cell modules in electroluminescence image. *Nondestruct. Test. Eval.* **2025**, *40*, 309–331. [\[CrossRef\]](#)
16. Gan, Y.; Ren, X.; Liu, H.; Chen, Y.; Lin, P. A novel lightweight detector for precise measurement of defects in photovoltaic panels in complex industrial scenarios. *Nondestruct. Test. Eval.* **2025**, 1–29. [\[CrossRef\]](#)
17. Luo, J.; Liu, Z.; Aleksandrovich Martynov, S.; Zhao, X.; Tan, J.; Lu, S. Detection algorithm for solar photovoltaic cell surface defects based on multi-scale edge information selection. *Nondestruct. Test. Eval.* **2025**, 1–46. [\[CrossRef\]](#)
18. Zhang, Y.; Zhang, X.; Tu, D. Unsupervised defect detection for solar photovoltaic cells based on convolutional autoencoder. *Nondestruct. Test. Eval.* **2025**, *40*, 6143–6159. [\[CrossRef\]](#)
19. 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, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
20. Sohan, M.; Sai Ram, T.; Rami Reddy, C.V. A review on Yolov8 and Its Advancements. In *International Conference on Data Intelligence and Cognitive Informatics*; Springer: Singapore, 2024; pp. 529–545.
21. Yang, L.; Zhang, R.; Li, L.S.; Simam, X. A simple, parameter-free attention module for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Honolulu, HI, USA, 23–29 July 2023; pp. 21–24.
22. Zheng, Z.; Wang, P.; Ren, D.; Liu, W.; Ye, R.; Hu, Q.; Zuo, W. Enhancing geometric factors in model learning and inference for object detection and instance segmentation. *IEEE Trans. Cybern.* **2021**, *52*, 8574–8586. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Tong, Z.; Chen, Y.; Xu, Z.; Yu, R. Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism. *arXiv* **2023**, arXiv:2301.10051.
24. Hou, Q.; Zhou, D.; Feng, J. Coordinate Attention for Efficient Mobile Network Design. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021.
25. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
26. Zheng, Z.; Wang, P.; Liu, W.; Li, J.; Ye, R.; Ren, D. Distance-IoU loss: Faster and better learning for bounding box regression. In Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; pp. 12993–13000.
27. Rezatofighi, H.; Tsoi, N.; Gwak, J.; Sadeghian, A.; Reid, I.; Savarese, S. Generalized intersection over union: A metric and a loss for bounding box regression. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 658–666.

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.