

Article

Detection of Small Targets in Photovoltaic Cell Defect Polarization Imaging Based on Improved YOLOv7

Haixia Wang¹, Fangbin Wang^{1,2,*}, Xue Gong^{1,2}, Darong Zhu^{1,2}, Ruinan Wang³ and Ping Wang¹

¹ School of Mechanical and Electrical Engineering, Anhui Jianzhu University, Hefei 230601, China; whx1229042022@163.com (H.W.); gongx@ahjzu.edu.cn (X.G.); gkj@ahjzu.edu.cn (D.Z.); pingwangliwei@ahjzu.edu.cn (P.W.)

² Polarized Light Imaging Detection Technology Anhui Provincial Key Laboratory, Hefei 230031, China

³ School of Mathematics and Science, Southeast University, Nanjing 210096, China; pfeng@seu.edu.cn

* Correspondence: wangfb@ahjzu.edu.cn

Abstract: A photovoltaic cell defect polarization imaging small target detection method based on improved YOLOv7 is proposed to address the problem of low detection accuracy caused by insufficient feature extraction ability in the process of small target defect detection. Firstly, polarization imaging technology is introduced, using polarization degree images as inputs to enhance the edge contour information of YOLOv7 for detecting small targets; then, the COT self-attention mechanism is added to reconstruct the SPPCSPC module to improve YOLOv7's ability to capture and fuse small target features in complex backgrounds; next, the normalized Wasserstein distance (NWD) is used to replace the traditional loss function based on intersection over union (IoU) metric, reducing the boundary offset between the prior box and the closest real target box in the prediction process of the object detection model and reducing the sensitivity of the YOLOv7 network to small object position deviations; finally, by constructing a shortwave infrared polarization imaging system to obtain polarization images of photovoltaic cells and detect small targets with scratch defects in photovoltaic cells, the applicability and effectiveness of the proposed method are verified. The results show that the proposed method has good recognition ability for small target defects in photovoltaic cells. By applying the constructed dataset, the detection accuracy reaches 98.08%, the recall rate reaches 95.06% and the mAP reaches 98.83%.

Keywords: photovoltaic cells; object detection; YOLOv7; polarization imaging



Citation: Wang, H.; Wang, F.; Gong, X.; Zhu, D.; Wang, R.; Wang, P. Detection of Small Targets in Photovoltaic Cell Defect Polarization Imaging Based on Improved YOLOv7. *Appl. Sci.* **2024**, *14*, 3899. <https://doi.org/10.3390/app14093899>

Academic Editor: Alejandro Pérez-Rodríguez

Received: 20 March 2024

Revised: 28 April 2024

Accepted: 29 April 2024

Published: 2 May 2024



Copyright: © 2024 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/>).

1. Introduction

Photovoltaic cells have high requirements for both conversion efficiency and working life, with even small defects having significant impacts. In addition, defects in these cells can generate excessive heat during the power generation process, potentially leading to cell module burnout due to overheating. This can result in fires within photovoltaic power plants, causing considerable economic losses [1]. An automatic defect detection method based on computer vision can identify and locate defects well, so it has a wide range of applications in the process of photovoltaic cell production and power plant operation and maintenance. Nevertheless, in practical applications, the accuracy of automatic detection is often limited by the small scale and subtle characteristics of defects. Therefore, how to achieve real-time and accurate detection of small defect targets is a practical problem that needs to be solved in the current operation and maintenance process of photovoltaic power stations, which has important practical significance for improving the performance of photovoltaic modules, extending the life of photovoltaic cells and reducing power generation costs [2].

Conventional photovoltaic cell defect detection techniques mainly include infrared thermal imaging, photoluminescence imaging and electroluminescence imaging. Among these, infrared thermal imaging technology [3] is a method of applying reverse bias to

photovoltaic modules to photograph infrared rays emitted by the surface of solar cells after local heating. However, this method suffers from low detection efficiency, ignores the distribution property of series resistance and incurs high costs. Photoluminescence imaging technology [4] is a non-contact and non-destructive imaging method that excites fluorescence from the silicon layer of photovoltaic modules under specific light sources, without the need for electricity. However, this method has a high economic cost [5]. Electroluminescence imaging [6] technology applies reverse voltage to both ends of the photovoltaic module, so that the silicon material of the battery will glow due to the current passing through. This method has high imaging quality and moderate cost and is currently the mainstream solar cell image acquisition method in the photovoltaic field. Nevertheless, the acquisition of electroluminescent images is susceptible to environmental stray light and camera imaging characteristics, leading to background interference, random noise, image blur and overall quality degradation [7]. Additionally, the subtle difference in luminous intensity between the defect edge and the normal part poses challenges in achieving satisfactory detection results.

Polarization is one of the essential characteristics of light radiation electromagnetic waves, less affected by the environment, so the use of polarization characteristics of imaging to obtain target information can improve the contrast between the target and the background and enhance the clarity of the target in the image. At present, polarization imaging target detection technology has been applied in many fields. Wang et al. [8] demonstrated through experiments that polarization imaging can emphasize the contour edges and texture details of photovoltaic cell defect images. Additionally, the Laplace pyramid is employed to decompose the captured light-intensity images and polarization degree images, fusing high- and low-frequency information. A defect detection method for crystalline silicon photovoltaic cells based on electroluminescence polarization image fusion is proposed, effectively highlighting the defect characteristics of photovoltaic cells. Zhu et al. [9] proposed an improved YOLOv5 method for detecting small targets in polarization imaging by inputting polarization degree images at the input end, reducing C3 modules and incorporating coordinate attention mechanisms. This method enhances the feature extraction ability and detection accuracy of small target images. Han et al. [10] proposed a target detection method with enhanced polarization information, which can improve the accuracy of common target recognition results in road scenes under haze weather conditions. Zhao et al. [11] proposed an object detection method based on local contrast of polarization weights, which can detect objects directly on infrared polarized mosaic images. Amer et al. [12] studied a method based on polarization imaging optical systems and a dark channel prior optimization version to improve the quality of underwater images collected in turbid environments, reducing the impact of diffusion on image acquisition and improving the image quality for automatic detection and recognition of underwater objects. Katletz et al. [13] designed a device based on circularly polarized terahertz waves and polarization-sensitive detectors, using the spectral information contained in broadband terahertz pulses, which allows for the study of birefringence and optic-axis orientation of glass fibers in polymers with a single measurement. Mazen et al. [14] proposed an electroluminescence defect detection method for PV cells based on improved YOLOv5 in order to maintain and inspect the cells of PV modules more efficiently and quickly, and the improved YOLOv5 algorithm achieved an average accuracy of 76.3%, which is 2.5% higher than the standard YOLOv5 algorithm.

With the rapid development of deep learning, convolutional neural networks (CNNs) [15] have received widespread attention in defect detection of photovoltaic modules due to their powerful feature extraction and data processing capabilities. In practical applications, this method is usually divided into two categories. One is a two-stage object detection algorithm based on candidate boxes represented by R-CNN networks, which has high detection accuracy but a slow running speed; another is a single-stage object detection algorithm represented by the YOLO model which has a fast running speed but slightly lower recognition accuracy than two-stage networks. Meng et al. [16] proposed a YOLO-PV

target detection algorithm by reducing the extraction of deep trunk information, increasing the extraction of low-level defect information and using a PAN network to fusion target features. The algorithm achieved 94.55% AP on the photovoltaic module EL image data set. Jamming speed exceeds 35 fps. Li et al. [17] introduced a bottleneck CSP module, added a tiny target detection prediction head in YOLOv5 to mitigate tiny defect missing, utilized Ghost convolution to improve the model inference speed and reduce the number of parameters and proposed a PV panel defect detection method based on the bottleneck CSP's Ghost convolution and the tiny target prediction head (GBH-YOLOv5) based on YOLOv5, which raises the mAP performance index to 27.8%.

In summary, the YOLO series network has shown good performance in defect detection, but there are still many difficulties in detecting small defect targets in photovoltaic cells, mainly manifesting as follows: (1) Traditional electroluminescence detection obtains grayscale images of photovoltaic cells, where small defect targets account for a small proportion of the entire image and the edges are blurred, making them susceptible to background interference. (2) With the continuous training and downsampling of the YOLOv7 network, many small defect features gradually disappear. (3) The bounding box loss of object detection algorithms is mostly based on the IoU metric, but it is not a good metric for small targets.

To address the aforementioned issues, this paper proposes a small target detection method for photovoltaic cell defect polarization imaging based on improved YOLOv7. This method can more accurately identify and detect small defects in photovoltaic cells. The main contributions of this article can be summarized as follows:

1. Introduce polarization imaging in electroluminescence to obtain polarization images of small defects in photovoltaic cells and analyze the polarization degree image as input to the YOLOv7 network. Polarization can remove certain background information interference and increase the contrast between defects and the background.
2. Add the contextual self-attention mechanism to reconstruct the relevant pyramid pooling structure module, increase the available target information in the image and improve the YOLOv7 network's ability to capture the features of small defects in photovoltaic cells.
3. The normalized Wasserstein distance is used to replace the traditional loss metric function based on the intersection of concatenation, to reduce the boundary offset between the a priori frame and the closest real target frame in the prediction process of the target detection model, to reduce the sensitivity of the YOLOv7 network to the positional deviation of the small defects of the PV cells and to increase the model's detection accuracy.

2. Relevant Theoretical Basis

2.1. Principles of Polarization Imaging in Electroluminescence

Crystal silicon solar photovoltaic cells typically exhibit an N+P homojunction structure, characterized by the addition of a heavily doped N-type semiconductor layer atop a P-type silicon substrate. In this configuration, the cell's reverse side functions as the positive electrode, while the front side serves as the negative electrode. The outermost layer of the cell is an anti-reflective coating, designed to minimize energy losses resulting from light reflection during operational conditions.

According to the composition of the photovoltaic cell depicted in Figure 1, it can be discerned as a PN junction structure. When a forward voltage bias is applied to photovoltaic cells, the electrons and holes in the P-N junction within the cell undergo radiative recombination, emitting near-infrared light. Should there be defects such as impurities, deposition or dislocations present in the silicon material, electrons and holes will preferentially recombine at these sites, consequently reducing or preventing near-infrared light emission. Similarly, if there are issues with the fabrication process of photovoltaic cells, such as silicon wafer fractures or other detrimental conditions, these may lead to problems with electrical

injection, subsequently diminishing the electroluminescence intensity of the affected areas. In the electroluminescence image, these areas will manifest as darker regions.

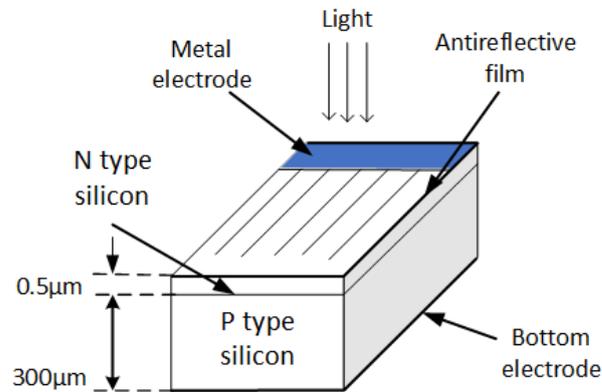


Figure 1. Structure of photovoltaic cells.

Light is a transverse wave whose vibration plane is perpendicular to its propagation direction. It exhibits several characteristics, including propagation direction, intensity and wavelength. Moreover, it possesses polarization characteristics. Polarized light can be categorized into fully polarized and partially polarized varieties, with fully polarized light further subdivided into circular and linear polarization types. Figure 2 shows the difference between unpolarized natural light and linearly polarized light: the light emanating from the source in the figure exhibits arbitrary vibration directions, thus classifying it as unpolarized light. Upon passage through a polarization filter, since only light oscillating along a specific direction is permitted to transmit, light with other vibration orientations is either absorbed or reflected. Consequently, the transmitted light becomes completely linearly polarized.

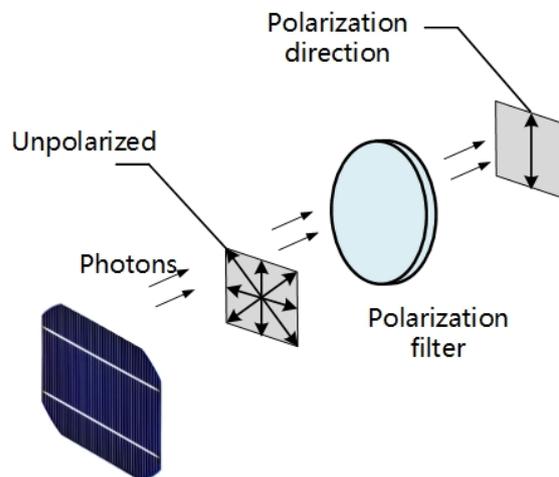


Figure 2. Polarized light imaging schematic diagram.

The Stokes vector is usually used to describe the polarization of light. It is assumed that the Stokes vector of the optical signal entering the imaging observation system with the polarization azimuth of 0°, 45°, 90° and 135° is S [18]. Then there is:

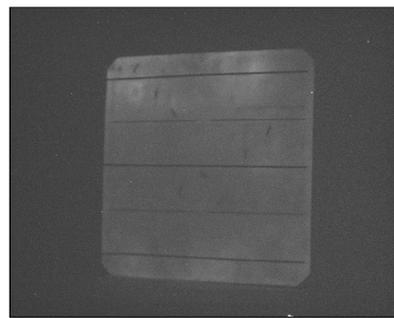
$$S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} I_0 + I_{45} \\ I_0 - I_{90} \\ I_{45} - I_{135} \\ 0 \end{pmatrix} = \begin{pmatrix} I_{90} + I_{135} \\ I_0 - I_{90} \\ I_{45} - I_{135} \\ 0 \end{pmatrix} \quad (1)$$

In the formula, S_0 represents the total intensity of light, S_1 represents the intensity of horizontal polarized light, S_2 represents the intensity of diagonal polarized light, and S_3 represents the intensity of circular polarized light. In general, the circular polarization component is very small and can be ignored. I_0 , I_{45} , I_{90} , and I_{135} represent the intensity of the light detected by a polarization observation system with a polarization azimuth of 0° , 45° , 90° , or 135° , respectively.

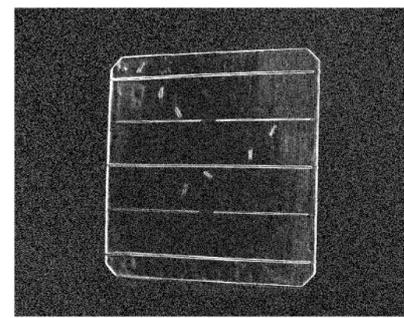
According to Formula (1), the polarization degree of the incident light can be determined as [18]:

$$DOP = \sqrt{S_1^2 + S_2^2} / S_0 \quad (2)$$

Figure 3 presents the intensity and polarization degree images obtained through electro-luminescence polarization imaging analysis of scratch defects on photovoltaic cells. As is evident from Figure 2, the surface of photovoltaic cells and the surface with scratch defects exhibit marked disparities in polarization images, primarily due to the fact that the surface of non-defective photovoltaic cells is relatively smooth, whereas the surface of scratched defective cells is comparatively rough. Hence, polarization imaging is employed to enhance the contrast of small defective targets within cell surface images, resulting in enhanced texture details and edge contours of the surface image of such small defect targets.



(a) Electroluminescence original picture



(b) Polarization picture

Figure 3. Schematic diagram of electroluminescence original image and polarization degree image.

2.2. YOLOv7 Algorithm

YOLOv7 [19] is the basic model in the YOLO series, which exceeds the speed and accuracy of most existing target detectors in the range of 5–160 frames/s [20]. The structure of YOLOv7 consists of four main parts: the input, the backbone feature extraction network, the enhanced feature extraction network and the head prediction portion.

In general, YOLOv7 initially resizes the input image to 640×640 , subsequently feeding it into the backbone network. The backbone of YOLOv7 primarily incorporates a CBS module, ELAN module, MPCConv module and SPPCSPC module. The CBS module comprises a Conv layer, BN layer and SiLU activation function. ELAN modules effectively facilitate learning and convergence by controlling gradient paths, whereas deeper networks employ MPCConv modules for downsampling. SPPCSPC utilizes both spatial pyramid pooling (SPP) and CSP structures to achieve richer gradient combinations. The enhanced feature extraction network implements multi-scale feature fusion and adaptive convolutional operations, enabling YOLOv7 to attain a significantly faster detection speed while maintaining high detection accuracy and meeting real-time requirements. Subsequently, the header network outputs three layers of feature maps of varying sizes, and the prediction results are ultimately generated via RepConv. The network structure of YOLOv7 is shown in Figure 4.

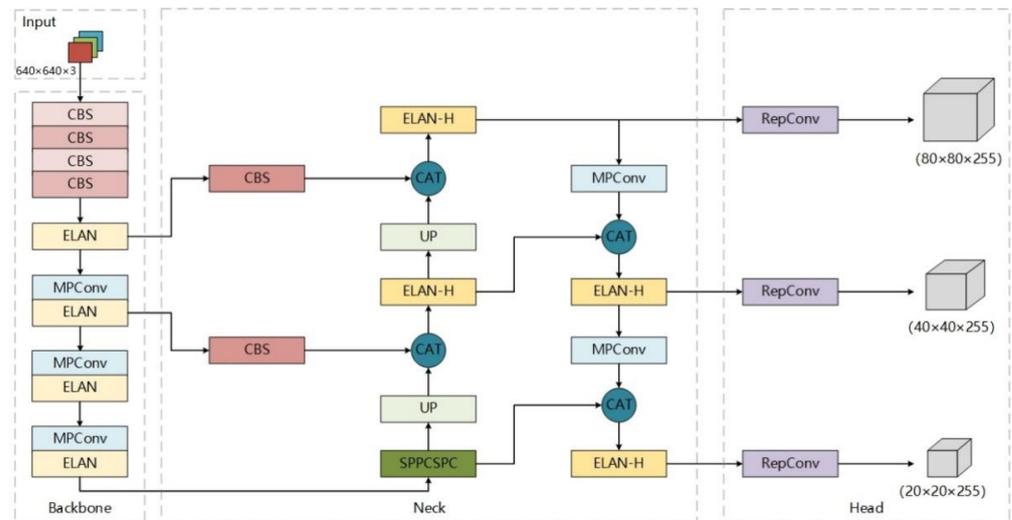


Figure 4. YOLOv7 structural block diagram.

3. Improvement of Defect Detection Algorithm

Based on YOLOv7's network, corresponding improvements have been made to the characteristics of small target polarization images. A contextual (COT) self-attention mechanism has been added to the SPPCSPC module, which can increase the information between the target and the background and replace the union intersection (IOU) metric with the normalized Wasserstein distance (NWD) metric to optimize the loss function and reduce the sensitivity of the YOLOv7 network to position deviations of small target objects.

3.1. COT Self Attention Mechanism

In polarization-degree images, extracting discriminative features proves challenging due to the small size of objects and the limited visualization information. In the YOLOv7 network, as convolutional layers deepen, the high contrast of small and medium targets in polarization degree images against the background tends to be diminished, further obscuring the boundary between the target and the background, thereby compromising the model's recognition accuracy and detection output. Hence, it is imperative to maximize the information dimension and depth accessible to the defect small targets. Typically, in acquired digital images, an inherent coexistence relationship exists between "target and scene" and "target and target" and leveraging this coexistence can significantly enhance the algorithm's detection performance for small defective targets [21]. From a research perspective [22], appropriately modeling contextual information can boost target detection performance, particularly when the network's depth increases, as it can more effectively demonstrate the enhancing effect of the attention mechanism. Grounded in this understanding, this paper incorporates the COT self-attention mechanism into YOLOv7 [23].

The COT structure is depicted in Figure 5. In contrast to a traditional self-attention mechanism, where all pairs of query key relationships are learned in isolation without considering the rich context between them, the COT structure overcomes this limitation. It efficiently leverages the contextual information between input keys to inform the learning of a dynamic attention matrix, thereby enhancing the visual representation with information beneficial to small target detection and yielding improved model performance.

In COT, set the input feature map $X \in \mathbb{R}^{H \times W \times C}$. The key, query and value are defined as $K = X$, $Q = X$ and $V = XW_v$. First use the $k \times k$ group convolution spatially for all adjacent keys within the $k \times k$ grid to represent each key in context. The learned context key $K^1 \in \mathbb{R}^{H \times W \times C}$ naturally reflects the static context information between local adjacent keys and K^1 is represented as the static context of input X . The context key K^1 and the query Q are cascaded and the attention matrix [23] is implemented

by two successive 1×1 convolution (W_θ with ReLU activation function and W_δ without activation function).

$$A = [K^1, Q] W_\theta W_\delta \tag{3}$$

For each header, the local attention matrix at each spatial location of A is learned based on query features and key features of the context, rather than isolated query keyword pairs. This approach enhances self-attention learning with the additional guidance of the mined static context K^1 . Next, the focus feature graph K^2 [23] is calculated by aggregating all values V based on the context's attention matrix A ,

$$K^2 = V \otimes A \tag{4}$$

Given the concern that feature graph K^2 captures dynamic feature interactions between inputs, K^2 is named the dynamic context of the inputs. The final output feature graph Y is the fusion of static K^1 and dynamic context K^2 .

COT can simultaneously capture the above two spatial context relationships in the input key, increase the effective information in the feature map and improve the effective extraction of small scratch features on the surface of the photovoltaic module. Therefore, in this paper, a COT self-attention mechanism is added to SPPCSPC module, which can improve accuracy and does not produce a large amount of calculation. The improved SPPCSPC module is shown in Figure 6.

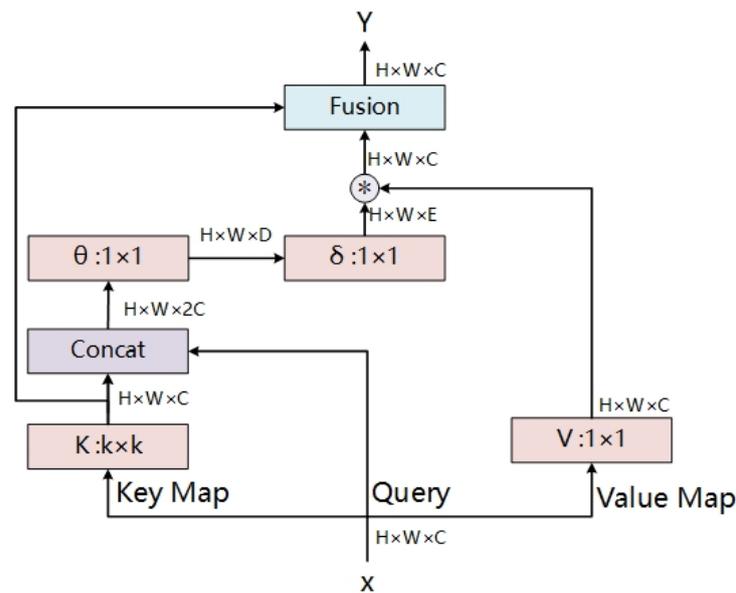


Figure 5. COT self-attention mechanism.

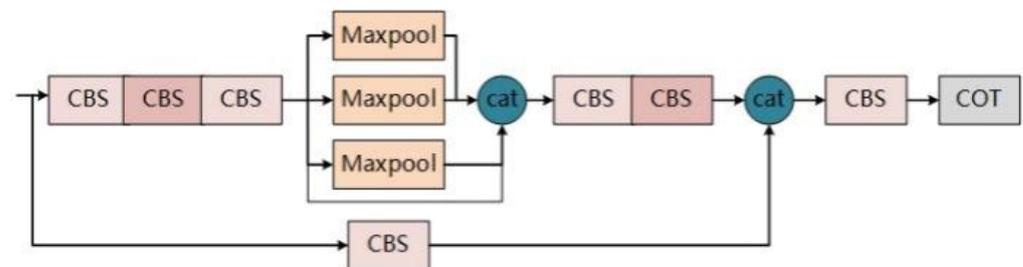


Figure 6. Improved SPPCSPC module.

3.2. Loss Function Based on NWD

Due to the limited area occupied by the small target in the polarization degree image, even a slight offset of one pixel in the bounding box prediction can lead to significant errors

in the prediction process. However, the YOLOv7 network lacks a bounding box regression function specifically designed for small targets, resulting in low regression accuracy when dealing with them. Therefore, we need to choose a loss function that is more tailored to small targets. Based on this, the regression loss of normalized Wasserstein distance (NWD) is chosen as the target loss function in YOLOv7.

The traditional loss function, which is based on the intersection over union (IoU) measure, has limited sensitivity to small targets due to the discreteness of boundary frame position changes. This sensitivity can lead to small position deviations, resulting in cases where the IoU between the true frame and any anchor falls below the minimum positive threshold. Consequently, a detection network trained on small targets lacks sufficient monitoring information, significantly reducing detection performance. On the other hand, NWD calculates the similarity between the small target prediction box and the ground truth box by normalizing the corresponding two-dimensional Gaussian distributions. This approach smooths out position deviations and can measure the similarity between boundary boxes that do not overlap or contain each other [24]. Furthermore, NWD is not sensitive to objects of different scales, which can improve the detection accuracy of the network model for small target objects.

In the bounding box, the target pixel and the background pixel are concentrated in the center and border of the bounding box respectively [25]. Therefore, modeling the bounding box as a two-dimensional (2D) Gaussian distribution, where the center pixel has the highest weight and the importance of pixels gradually decreases from the center towards the boundary, can better describe the weights of different pixels in the bounding box. Specifically, for bounding boxes $R = (c_x, c_y, w, h)$, where (c_x, c_y) , w and h represent center coordinates, width and height, respectively. The equation of its intrinsic ellipse can be expressed as

$$\frac{(x - \mu_x)^2}{\sigma_x^2} + \frac{(y - \mu_y)^2}{\sigma_y^2} = 1 \tag{5}$$

where (μ_x, μ_y) are the central coordinates of the ellipse and σ_x and σ_y are the lengths of the semi-axes along the x and y axes. Therefore, $\mu_x = c_x, \mu_y = c_y, \sigma_x = w/2, \sigma_y = h/2$.

The probability density function of a 2D Gaussian distribution is given by the following equation:

$$f(x|\mu, \Sigma) = \frac{\exp(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu))}{2\pi|\Sigma|^{\frac{1}{2}}} \tag{6}$$

The similarity between boundary boxes A and B can be converted to the distribution distance between two Gaussian distributions, which is calculated using the Wasserstein distance from optimal transport theory. For two two-dimensional Gaussian distributions $\mu_1 = N(m_1, \Sigma_1)$ and $\mu_2 = N(m_2, \Sigma_2)$, the second-order Wasserstein distance between μ_1 and μ_2 is defined as [26]:

$$W_2^2(\mu_1, \mu_2) = \|m_1 - m_2\|_2^2 + \text{Tr}(\Sigma_1 + \Sigma_2 - 2(\Sigma_1^{1/2}\Sigma_2\Sigma_1^{1/2})^{1/2}) \tag{7}$$

which can be simplified as:

$$W_2^2(\mu_1, \mu_2) = \|m_1 - m_2\|_2^2 + \|\Sigma_1^{1/2} - \Sigma_2^{1/2}\|_F^2 \tag{8}$$

Based on the Gaussian distributions N_a and N_b modeled in boundary boxes $A = (c_{x_a}, c_{y_a}, w_a, h_a)$ and $B = (c_{x_b}, c_{y_b}, w_b, h_b)$, Equation (8) can be further simplified as

$$W_2^2(N_a, N_b) = \left\| \left(\begin{bmatrix} c_{x_a}, c_{y_a}, \frac{w_a}{2}, \frac{h_a}{2} \end{bmatrix}^T, \begin{bmatrix} c_{x_b}, c_{y_b}, \frac{w_b}{2}, \frac{h_b}{2} \end{bmatrix}^T \right) \right\|_2^2 \tag{9}$$

However, $W_2^2(N_a, N_b)$ is a distance measure and cannot be used directly as a similarity measure (i.e., a value between 0 and 1 as an IoU). Therefore, we use its exponential form to normalize and obtain a new metric called normalized Wasserstein distance (NWD):

$$NWD(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{C}\right) \quad (10)$$

where C is a constant that is closely related to the data set.

We designed the NWD metric as a loss function in the following way:

$$Loss_{NWD} = 1 - NWD(N_p, N_g) \quad (11)$$

where N_p is the Gaussian distribution model of the prediction box and N_g is the Gaussian distribution model of the truth box.

NWD can be used as a better metric to measure the similarity between two bounding boxes, overcoming the sensitivity of IOU to the position deviation of small target objects and becoming embedded in the target loss function to improve the detection of small targets by the YOLOv7 network.

4. Experiment and Result Analysis

4.1. Experimental Platform and Dataset Construction

A time-multiplexed polarization imaging system was selected to capture the polarization images of small scratches on the surface of photovoltaic cells, as shown in Figure 7. Images of polarized light intensity in different directions were obtained by placing a rotatable polarizer in front of the imaging detector. The whole system is mainly composed of three parts: a shortwave camera, DC power supply, polarizer turntable and laptop computer with polarizing image acquisition software. The shortwave infrared camera model is Owl 640 S, featuring a response wavelength ranging from 900 nm to 1700 nm and an internal InGaAs sensor. The polarizer is a precision line grid polarizer, model 20LP-NIR, with an extinction ratio of 1000:1. The excitation power supply is a Goodyear 3030DD DC power supply, capable of providing a constant current output within a voltage range of 0–30 V and a current range of 0–3 A. The experimental object is a monocrystalline silicon solar cell sheet, measuring 156.6 mm × 156.6 mm with a thickness of approximately 0.1 mm. The scratches on the cell have a maximum length of 9 mm and a width of 1 mm. The positive and negative poles of the cell are led out with leads to facilitate connection to the DC regulated power supply.

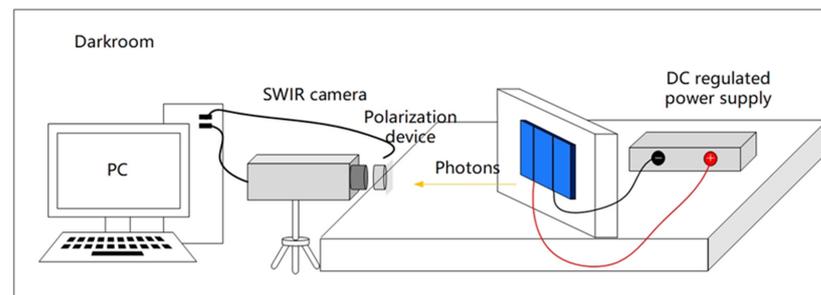


Figure 7. Short wave polarized infrared shooting device diagram.

The experiment was conducted in a dark room, with the photovoltaic cell positioned perpendicular to the surface of the mobile platform. The distance between the camera lens and the center of the photovoltaic cell was maintained at 60 cm, and the DC power output was set at 1V-3A to make the photovoltaic cell positive. Multiple images of the photovoltaic cell were captured at different observation angles, corresponding to polarization azimuths of 0°, 45°, 90° and 120°. The Stokes parameters and polarization degree images were

obtained by means of Equations (1) and (2). The DOP images obtained by analysis are marked with marking software, and the marking information is exported in txt form.

The key parameters of the deep learning platform in this experiment are as follows: Operating system: Ubuntu 18.04; graphics card model: NVIDIA RTX A 2080Ti with 11 GB memory. CPU model: Intel(R) Xeon(R) Platinum 8255C CPU@2.50 GHz. Deep learning framework Pytorch1.7.0; CUDA version 11.0. The defect data set was randomly divided into training set, test set and verification set according to the ratio of 7:2:1, with 309, 87 and 35 pictures, respectively.

4.2. Definition of Evaluation Indicators

In the stage of data analysis, six indexes were used to evaluate the model effect, including accuracy rate (P), recall rate (R), $mAP_{0.5}$, $mAP_{0.5:0.95}$, training time and single image test time. Accuracy rate, recall rate and mAP (mean average precision) are all metrics that assess the accuracy of the model's detections and their formulas are shown below. Training time serves as a measure of the detection model's training efficiency, while single image test time represents a measure of the speed of image processing by the object detection network.

$$P = \frac{TP}{TP + FP} \quad (12)$$

$$R = \frac{TP}{TP + FN} \quad (13)$$

$$AP = \int_0^1 P(R) \quad (14)$$

$$mAP = \frac{1}{c} \sum_{j=1}^c AP_j \quad (15)$$

Here are the definitions: P represents accuracy, R stands for recall rate, TP is the number of true positives (positive samples correctly determined), FP is the number of false positives (samples incorrectly judged as positive), FN is the number of false negatives (samples incorrectly judged as negative), AP represents the average accuracy across all classes, and c denotes the total number of classifications.

4.3. Result Analysis

In the training process, Figure 8 compares the training loss and $mAP_{0.5}$ curves of both the original and improved YOLOv7 algorithms. The horizontal axis represents the training times, while the vertical axis in Figure 8a corresponds to the loss value and in Figure 8b to the $mAP_{0.5}$ value. Figure 8a demonstrates that the improved YOLOv7 algorithm converges faster and achieves a lower final loss value compared to the original algorithm. In Figure 8b, it is evident that the improved YOLOv7 algorithm exhibits less jitter and more stable mAP performance, resulting in a higher final value.

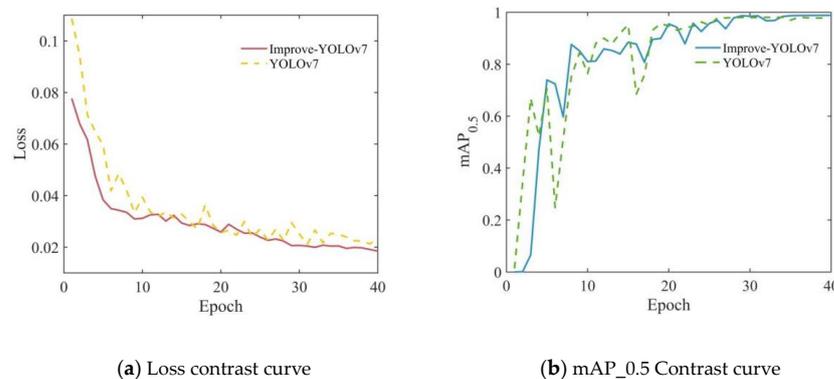


Figure 8. Training loss and $mAP_{0.5}$ comparison curve.

4.3.1. Contrast Experiment

In order to objectively demonstrate the improved effect of the proposed algorithm, the improved model was compared and analyzed with official models such as YOLOv7, YOLOv7-Tiny and YOLOv7x. The analyzed results of various indicators are shown in Table 1, and the parameters used by the model were the same. The experimental results show that, compared with the original YOLOv7 algorithm, the improved YOLOv7 algorithm sacrifices a detection speed of 0.001s but achieves the best performance in terms of accuracy, recall rate and mAP scores, including mAP_0.5 and mAP_0.5:0.95. It also outperforms other official YOLOv7 models across all four metrics. This demonstrates that the model exhibits excellent performance in terms of mean average accuracy and detection speed after loss function optimization and attention mechanism feature enhancement.

Table 1. Performance comparison of various models.

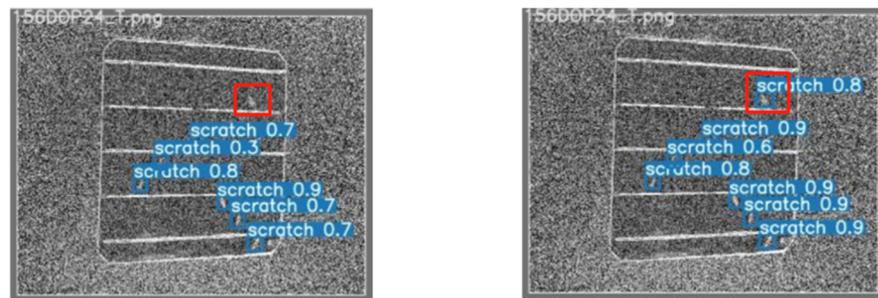
Model	Accuracy Rate/%	Recall Rate/%	mAP_0.5/%	mAP_0.5:0.95/%	Training Time/h	Test Time/s
YOLOv7	95.54	93.21	97.82	49	0.116	0.106
YOLOv7-Tiny	92.02	93.21	95.5	43.14	0.079	0.056
YOLOv7x	88.68	91.9	94.31	42.78	0.170	0.110
Improvement YOLOv7	98.08	95.06	98.83	50.34	0.117	0.107

4.3.2. Test Results

In order to verify the performance of the proposed algorithm in practical applications, pictures of three scenarios were extracted from the data test set and the proposed algorithm and the original YOLOv7 algorithm were respectively used for testing, as shown in Figure 9.

The results of Figure 9 show that in group (a), there are noise interference points in the background of photovoltaic cells; both models show good adaptability and no false detection occurs, but the original YOLOv7 algorithm fails to detect defect points and the proposed algorithm has higher confidence. In the (b) group of pictures, the grayscale between the photovoltaic cell defects and the background is relatively close and the original YOLOv7 algorithm also fails to detect points, which shows that the detection effect of the proposed algorithm on small targets has been improved. The scenario of group (c) is more complex, and small targets overlap. The original YOLOv7 detected one defect, but the algorithm in this paper could identify two defects, indicating that the model has strong recognition ability.

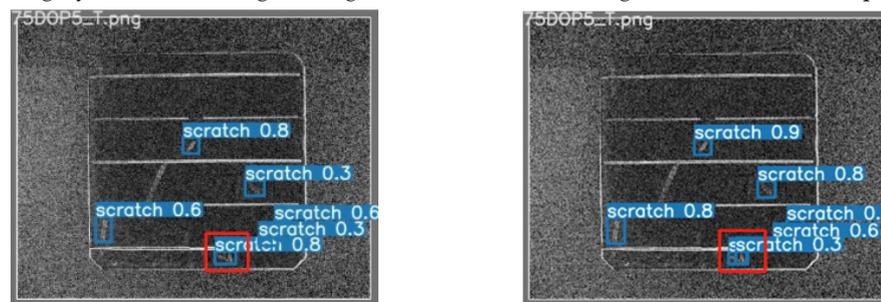
Figure 10 is a comparison between the middle feature map of the YOLOv7 model and the improved model. It can be seen from the figure that the edge outline of defects in the feature map becomes clearer after the COT self-attention mechanism module is added. The contrast between the defect and the background photovoltaic cell in the feature map is obviously improved.



(a) Comparison of background noise interference image detection effect.

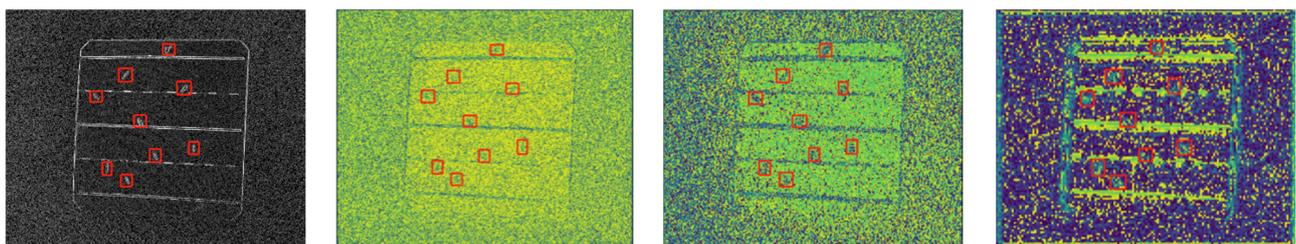


(b) The gray value of the target background is close to the image detection effect comparison.

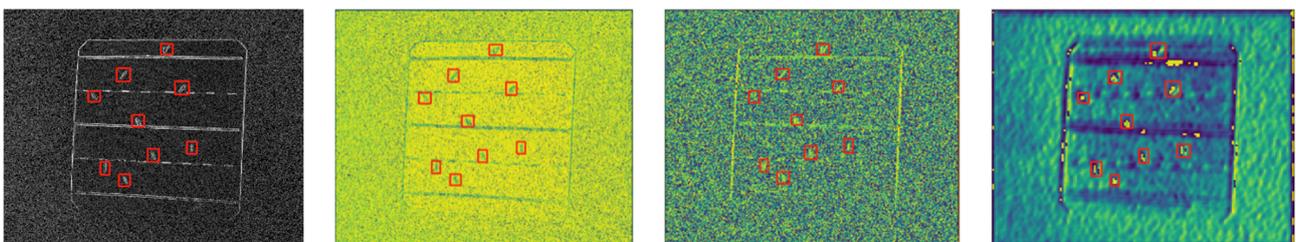


(c) Comparison of detection results of small target overlapping images.

Figure 9. Picture test results. The red boxes in the figure indicate the difference between the detection results of the original YOLOv7 algorithm and the algorithm in this paper for the same defect case.



(a) Intermediate feature map of the original algorithm.



(b) The intermediate feature map of the improved algorithm.

Figure 10. Comparison results of intermediate feature maps. The red boxes in the figure indicate the location of the defect.

5. Conclusions

This paper proposes a small target detection method based on improved YOLOv7 polarization imaging of photovoltaic cell defects. First, the polarization camera is used to collect small defects of photovoltaic modules and the analyzed DOP image information is used as network input to make the edge of the small target prominent, increase the available information of images and increase the detection accuracy of the model. Secondly, a COT attention mechanism module is added to increase the accessible information of small targets by using the context of small targets and improve the detection accuracy of small targets. Finally, by improving the target loss function and replacing the IOU measure with the NWD measure, the sensitivity of IOU to the position deviation of small target objects can be overcome, the prior box regression effect can be improved, and the detection accuracy of small target can be further improved. Therefore, the mAP has obvious improvements.

The experimental and research results show that the proposed method is 1.01 percentage points higher than the mAP of the original YOLOv7 algorithm. Subsequent research work will involve more comparative experiments in the field of PV cell defect detection and verify the applicability of the method proposed in this paper under different types of defect conditions. Attention will also be paid to the lightweighting of the model so that it can be run in mobile devices while maintaining accuracy.

Author Contributions: Conceptualization, H.W. and F.W.; methodology, H.W. and D.Z.; software, H.W.; validation, X.G. and P.W.; investigation, H.W. and R.W.; data curation, X.G.; writing—original draft preparation, H.W.; writing—review and editing, H.W., F.W. and D.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Anhui Natural Science Foundation (2008085UD09); Anhui University Collaborative Innovation Project (GXXT-2021-010); Anhui Construction Plan Project (2022-YF016, 2022-YF065, 2023-YF050); Anhui Province Higher Education Science Research Project (2022AH040044); Anhui Province University Outstanding Youth Research Project (2022AH020025).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

Acknowledgments: We would like to acknowledge the support from Anhui Natural Science Foundation, Anhui Provincial Department of Education, Anhui Provincial Department of Housing and Urban Rural Development.

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

References

1. Guerriero, P.; Daliento, S. Toward a hot spot free PV module. *IEEE J. Photovolt.* **2019**, *9*, 796–802. [[CrossRef](#)]
2. 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.
3. Frühauf, F.; Breitenstein, O. DLIT-versus ILIT-based efficiency imaging of solar cells. *Sol. Energy Mater. Sol. Cells* **2017**, *169*, 195–202. [[CrossRef](#)]
4. Trupke, T.; Mitchell, B.; Weber, J.W.; McMillan, W.; Bardos, R.A.; Kroeze, R. Photoluminescence imaging for photovoltaic applications. *Energy Procedia* **2012**, *15*, 135–146. [[CrossRef](#)]
5. Trupke, T.; Bardos, R.A.; Schubert, M.C.; Warta, W. Photoluminescence imaging of silicon wafers. *Appl. Phys. Lett.* **2006**, *89*, 044107. [[CrossRef](#)]
6. Kim, H.; Kim, J.Y.; Park, S.H.; Lee, K.; Jin, Y.; Kim, J.; Suh, H. Electroluminescence in polymer-fullerene photovoltaic cells. *Appl. Phys. Lett.* **2005**, *86*, 183502. [[CrossRef](#)]
7. Deitsch, S.; Buerhop-Lutz, C.; Sovetkin, E.; Steland, A.; Maier, A.; Gallwitz, F.; Riess, C. Segmentation of photovoltaic module cells in uncalibrated electroluminescence images. *Mach. Vis. Appl.* **2021**, *32*, 84. [[CrossRef](#)]
8. Wang, F.; Zhang, Y.; Wang, F.; Zhu, D. Polarization image fusion and defect detection of photovoltaic cell electroluminescence. *Electron. Meas. Technol.* **2022**, *45*, 143–149.

9. Zhu, D.; Wang, D.; Wang, F.; Gong, X.; Yang, G.; Yan, R. Improved YOLOv5 method for polarization imaging small target detection. *J. Chang. Norm. Univ.* **2023**, *42*, 58–66.
10. Han, H.; He, Y.; Du, X.; Wang, H. Multi object detection based on polarization information image enhancement. *J. Electron. Meas. Instrum.* **2023**, *37*, 29–38.
11. Zhao, Y.; Zhang, J.; Qiao, X. Object detection based on polarization weighted local contrast. *J. Infrared Millim. Waves* **2023**, *42*, 102–110.
12. Amer, K.O.; Elbouz, M.; Alfalou, A.; Brosseau, C.; Hajjami, J. Enhancing underwater optical imaging by using a low-pass polarization filter. *Opt. Express* **2019**, *27*, 621–643. [[CrossRef](#)] [[PubMed](#)]
13. Katletz, S.; Pflieger, M.; Pühringer, H.; Mikulics, M.; Vieweg, N.; Peters, O.; Scherger, B.; Scheller, M.; Koch, M.; Wiesauer, K. Polarization sensitive terahertz imaging: Detection of birefringence and optical axis. *Opt. Express* **2012**, *20*, 23025–23035. [[CrossRef](#)] [[PubMed](#)]
14. Mazen, F.M.A.; Seoud, R.A.A.; Shaker, Y. Deep learning for automatic defect detection in PV modules using electroluminescence images. *IEEE Access* **2023**, *11*, 57783–57795. [[CrossRef](#)]
15. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaria, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [[CrossRef](#)] [[PubMed](#)]
16. Meng, Z.; Xu, S.; Wang, L.; Gong, Y.; Zhang, X.; Zhao, Y. Defect object detection algorithm for electroluminescence image defects of photovoltaic modules based on deep learning. *Energy Sci. Eng.* **2022**, *10*, 800–813. [[CrossRef](#)]
17. Li, L.; Wang, Z.; Zhang, T. Gbh-yolov5: Ghost convolution with bottleneck csp and tiny target prediction head incorporating yolov5 for pv panel defect detection. *Electronics* **2023**, *12*, 561. [[CrossRef](#)]
18. Zhu, D.; Wang, D.; Wang, F.; Gong, X.; Yang, G.; Yan, R. Experimental study of the effect of fallen leaf shading on the polarization characteristics of photovoltaic modules. *SN Appl. Sci.* **2023**, *5*, 165. [[CrossRef](#)]
19. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023.
20. Qi, L.; Gao, J. Small object detection based on improved YOLOv7. *Comput. Eng.* **2023**, *49*, 41–48.
21. Gao, X.; Mo, M.; Wang, H.; Leng, J. Progress in Small Object Detection Research. *Data Collect. Process.* **2021**, *36*, 391–417. [[CrossRef](#)]
22. Oliva, A.; Torralba, A. The role of context in object recognition. *Trends Cogn. Sci.* **2007**, *11*, 520–527. [[CrossRef](#)]
23. Li, Y.; Yao, T.; Pan, Y.; Mei, T. Contextual Transformer Networks for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 1489–1500. [[CrossRef](#)] [[PubMed](#)]
24. Wang, J.; Xu, C.; Yang, W.; Yu, L. A Normalized Gaussian Wasserstein Distance for Tiny Object Detection. *arXiv* **2021**, arXiv:2110.13389.
25. Wang, J.; Yang, W.; Li, H.-C.; Zhang, H.; Xia, G.-S. Learning center probability map for detecting objects in aerial images. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 4307–4323. [[CrossRef](#)]
26. Panaretos, V.M.; Zemel, Y. Statistical aspects of Wasserstein distances. *Annu. Rev. Stat. Its Appl.* **2019**, *6*, 405–431. [[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.