

Article

YOLO v7-ECA-PConv-NWD Detects Defective Insulators on Transmission Lines

Jianrui Zhang ^{1,*}, Xia Wei ¹, Linxuan Zhang ², Libin Yu ³, Yannan Chen ¹  and Meiqi Tu ²

¹ School of Electrical Engineering, Xinjiang University, Urumqi 830017, China; xjuweix@xju.edu.cn (X.W.); xjuchenyin163@163.com (Y.C.)

² National Computer Integrated Manufacturing System Engineering Research Center, Tsinghua University, Beijing 100084, China; lxzhang@tsinghua.edu.cn (L.Z.); tmq21@mails.tsinghua.edu.cn (M.T.)

³ Urumqi Power Supply Section of China Railway Urumqi Bureau Group Co., Urumqi 830017, China; 13565849451@139.com

* Correspondence: 107552103980@stu.xju.edu.cn

Abstract: This paper proposes an enhanced YOLO v7-based method for detecting insulator defects in transmission lines, addressing the challenges of low accuracy and high leakage rates caused by complex backgrounds and electric poles alongside varying sizes of insulator targets in the image. Firstly, to address the issue of background interference and improve the importance of insulator features, a lightweight attention mechanism named Efficient Channel Attention (ECA) was introduced. With the incorporation of ECA, this model could effectively suppress background noise and provide more focus to insulator regions, thus enhancing its ability to detect insulator defects accurately. Secondly, a partial convolution (PConv) approach was employed in the backbone network instead of conventional convolution, which learned some important channels. This substitution improved both the network model's accuracy and the training speed. Finally, the Normalized Wasserstein Distance (NWD) prevented insulator features from being lost during pre-feature extraction, which reduced the leakage rate and improved the detection accuracy of small target insulators and defective insulators. The experimental results demonstrate that the improved YOLO v7 network model achieved an average detection accuracy (mAP) of 98.1%, recall of 93.7%, and precision of 96.8% on the TISLTR dataset. On the FISLTR dataset, the average detection accuracy (mAP) for flashover insulators was 93%, with a recall of 92.3% and precision of 87.1%. The average detection accuracy (mAP) for broken insulators was 92.2%, with a recall of 90.3% and a precision of 95.2%. These metrics demonstrate significant improvements in both datasets, highlighting the proposed algorithms' strong generalization capability and practicable potential to detect insulator targets.

Keywords: UAV Patrol Image; insulator identification; object detection; deep learning



Citation: Zhang, J.; Wei, X.; Zhang, L.; Yu, L.; Chen, Y.; Tu, M. YOLO v7-ECA-PConv-NWD Detects Defective Insulators on Transmission Lines. *Electronics* **2023**, *12*, 3969. <https://doi.org/10.3390/electronics12183969>

Academic Editor: Nurul I. Sarkar

Received: 21 August 2023

Revised: 11 September 2023

Accepted: 18 September 2023

Published: 20 September 2023



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

1. Introduction

In recent years, China's electric power industry has witnessed remarkable growth in tandem with the country's economic development. According to statistics, China's industrial electricity consumption is set to reach 5509 billion kWh in 2021, an increase of 9.1% year on year, while urban and rural residential electricity consumption could reach 1174.3 billion kWh, an increase of 7.3% year on year [1]. However, the manual inspection of transmission lines has become increasingly burdensome and cannot fully guarantee the normal operation of high-voltage equipment. As a result, frequent faults in transmission lines have emerged, necessitating higher requirements for the intelligent maintenance of power systems [2].

Insulators, as a special isolation control, can be used to fix the conductive body to ensure the smooth transmission of electricity. In the overhead, high-voltage transmission lines play an important role. Due to their exposure to outdoor environments and susceptibility to adverse weather conditions, insulators are prone to damage and require regular

inspection and maintenance. The inspection methods involved are relatively advanced. One effective method is the automatic detection of insulator defects, which can enhance productivity, protect staff from potential hazards, and mitigate safety risks. Therefore, the automatic defect detection of insulators holds great significance for improving maintenance efficiency [3].

Both domestically and internationally, the current approaches to insulator defect detection in transmission lines can be broadly classified into traditional image processing techniques, machine learning algorithms, and deep learning methods.

Traditional image processing methods, such as filtering [4], edge detection [5], and morphological processing [6], have been widely used. For example, Chen Guocui et al. [7] applied an improved fast-guided filtering algorithm to filter insulator images, effectively removing noise while preserving the edge detail features of the insulators. Zhao Le et al. [3] utilized an edge detection algorithm with a noise suppression module for power line feature extraction. Mei Xin et al. [8] employed morphological processing to optimize images of composite insulator surfaces covered with water, mitigating the influence of lighting conditions. These traditional methods are simple and easy to use, with high computational efficiency and clear mathematical models and algorithmic formulas. However, they are also susceptible to interference from background pixels and external noise and lack robustness and immunity to interference. Machine learning methods, on the other hand, leverage spatial and color feature information for insulator detection. Wu Yang et al. [9] applied the AdaBoost algorithm to insulator detection and recognition, demonstrating good robustness and laying the foundation for subsequent insulator fault diagnosis. Huang Huihang et al. [10] integrated machine learning modules into an insulator anomaly detection system, enabling the automatic detection of insulator anomalies.

Compared with traditional image processing methods, machine learning methods are more robust, can automatically analyze large amounts of data to extract valuable information, have some noise processing ability, and have some insulator defect detection capabilities. However, these methods are more suitable for offline data analysis and may require human intervention to achieve better detection results. In addition, machine learning methods have limited adaptability and generalizability in complex environments. In conclusion, both traditional image processing methods and machine learning methods have limitations when effectively solving complex problems and might not be able to meet the practical requirements of detection accuracy and speed.

With the continuous advancements in hardware technology, new deep learning detection methods have emerged, addressing challenges that traditional image processing methods and machine learning approaches struggle with. These deep learning methods can be generally classified into two-stage detection methods and single-stage detection methods.

Two-stage detection methods utilize a two-step process to detect targets [11]. Firstly, they generate candidate regions that potentially contain the insulators being detected. Then, each candidate region is classified and refined for recognition and localization. Some popular two-stage detection methods include R-CNN [12] and Faster R-CNN [13]. For instance, Zheng Ruojun [14] employed cropped R-CNN to extract insulator features and performed validation on Raspberry Pi, eliminating the need to transfer images to remote GPU servers for processing. Yi Jiyu et al. [15] incorporated multi-scale images and introduced an adversary generation strategy to enhance the accuracy of Faster RCNN when detecting blocked insulators. Tian Zijian et al. [16] proposed a two-stage target enhancement network specifically designed for low illumination environments, effectively detecting insulator faults.

The two-stage detection method offers greater accuracy when locating insulators and achieves higher overall detection accuracy, particularly on large-scale datasets. However, this method is more complex, demands more computational resources, and poses challenges in deployment on mobile devices.

Single-stage detection methods offer faster inference speed and significant engineering application values than two-stage detection methods. Some classic single-stage detection methods include the YOLO [17–21] series and the SSD [22] series.

For instance, Zhu Youchan et al. [23] utilized Darknet-53 as the feature extraction network in YOLOv3 to successfully detect and identify normal insulators. Song Libo et al. [24] employed Resblock-D+CSPDarknet53-tiny as the backbone network for YOLOv4-tiny and achieved successful deployment on Jetson NANO, showcasing its engineering usability. Wang Jianye [25] developed a lightweight multi-scale feature fusion SSD model to detect insulator self-detonation faults.

Compared to the two-stage detection method, the single-stage detection method outputs the defect detection results directly from the input image without multiple processing stages. Compared with the traditional multi-stage pipeline approach, it simplifies the complexity of the whole inspection system and improves the efficiency of the model. It facilitates insulator defect detection and model deployment. However, single-stage detection methods are more suitable for ordinary image comparisons and might struggle with high-resolution image detection. Insulator images often have complex backgrounds, variable sizes, and the insulator defect area is relatively small, which can lead to suboptimal detection results using single-stage methods.

Therefore, this paper presents a novel approach for detecting defects in transmission line insulators using an enhanced version of YOLO v7. This method aims to achieve the efficient and accurate detection of insulator faults while maintaining a high detection speed. The major contributions of this paper are outlined as follows:

- (1) Data Enhancement: The standard insulator dataset TISLTR and the high-resolution tiny target faulty insulator dataset FISLTR are enhanced through various techniques, including image enhancement, flipping, cropping, blurring, and random transformations. These techniques enhance the datasets by increasing their diversity and quality.
- (2) Attention Mechanism: To address the challenge of varying insulator sizes and occlusion caused by transmission line towers, the ECA is incorporated into the backbone of YOLO v7. This mechanism dynamically learns the significance of various channels in the input image, effectively reducing the impact of pole occlusion and enhancing the detection algorithm's accuracy.
- (3) Partial Convolution: The YOLO v7 network model has a complex network structure and many computational parameters, resulting in slower performance. To tackle this issue, partial convolution (PConv) is introduced as a replacement for traditional convolution in the YOLO v7 network model. PConv ensures both efficiency and detection accuracy while reducing computational parameters.
- (4) Normalized weighted detection: Due to the complex background of insulator images and the presence of small and dense insulators in datasets, insulator features can easily be lost during feature extraction, leading to missed detections. The YOLO v7 network model uses the Normalized Weighted Distance (NWD) metric instead of the traditional Intersection over Union (IoU) for target detection, which reduces the sensitivity of the IoU to the positional deviation of insulators with small targets, thus achieving the effective detection of insulators with small targets.

2. Enhanced YOLO v7 Defect Detection Method

2.1. Overview of the YOLO v7 Algorithm

The YOLO v7 network model was proposed by Chien-Yao et al. [26] in 2022 and compared to the previous YOLO v3, YOLO v4, YOLO v5, and YOLOX network models; it surpassed the previous YOLO series in terms of accuracy and detection precision in the range of 5 FPS to 160 FPS. The YOLO v7 network model mainly contains Input, Backbone, and Head. Input is the input image, Backbone is the backbone network that extracts the input image features, and Head is the detection head that predicts the type and location of the target, still using the YOLO v5 network architecture model.

This paper describes the structure of the improved YOLO v7 network model, as shown in Figure 1. First, the attention mechanism ECA is introduced in the Backbone of YOLO v7. It can adaptively learn the features of different channels to improve the accuracy of the detection algorithm. Second, the traditional convolution in the Backbone of YOLO v7 is replaced with PConv to reduce the computational parameters and ensure the efficiency of the network model. Finally, to reduce the sensitivity of *IoU* to the positional deviation of small target insulators, the target detection intersection and fusion ratio *IoU* in YOLO v7 is replaced with NWD, which reduces the leakage detection rate of small target insulators.

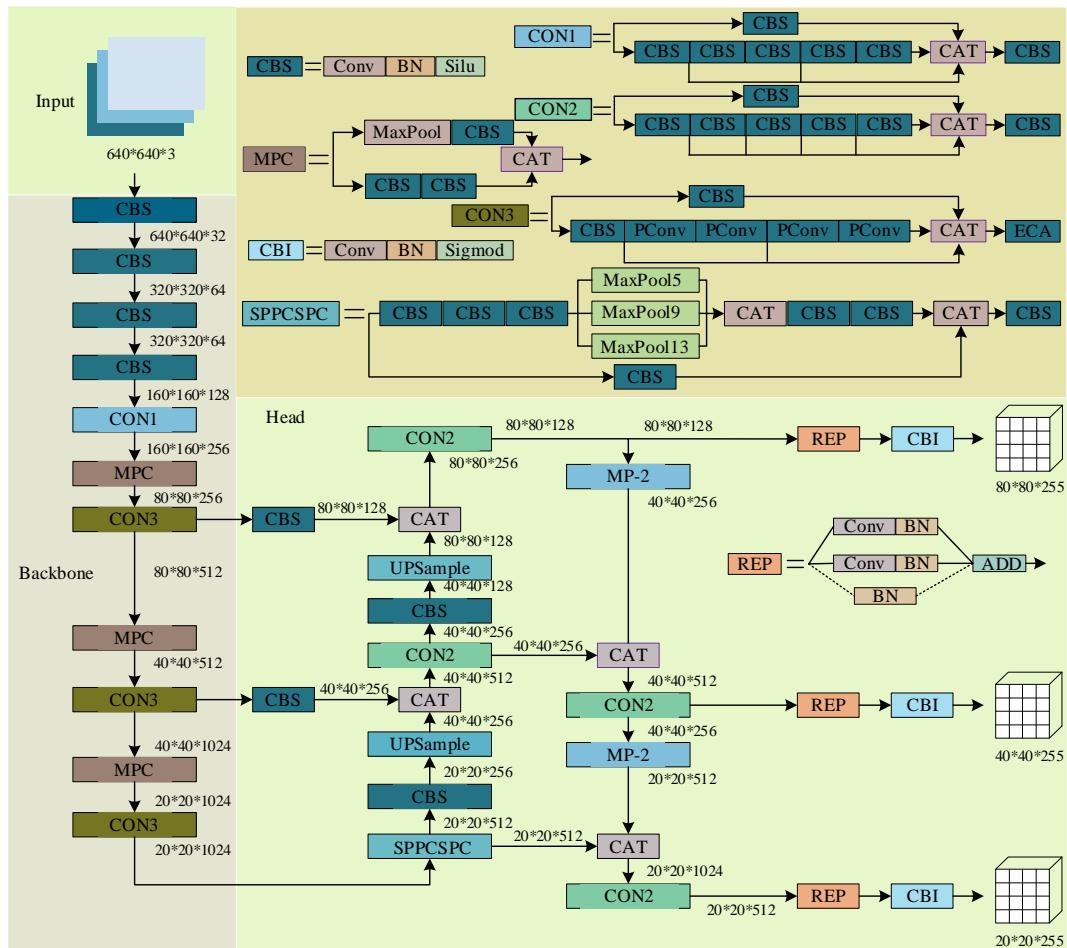


Figure 1. The structure diagram of the improved YOLO v7 network model.

2.2. Insulator Detection Based on the Attention Mechanism ECA

Transmission line insulator detection targets are of different sizes, the broken insulator site area is very small, and the detection background is complex and variable, while the original YOLO v7 network model tends to ignore the feature information of small targets during feature extraction, resulting in a lower detection accuracy during insulator target detection. Therefore, the channel attention module (Efficient Channel Attention, ECA) is added to the YOLO v7 backbone network Backbone [27], which means that YOLO v7 is based on the Efficient Aggregation Network E-ELAN. The channels that are effective in isolator feature extraction are given higher weights to suppress irrelevant background features and reduce the influence of background noise on target detection; therefore, the network model is able to judge the position and size of isolators.

In addition, ECA replaces the fully connected layer with a 1×1 convolutional layer on top of SE, avoiding dimensionality reduction and allowing the network model to achieve better results with fewer parameters. ECA Channel Attention the specific structure of the module is shown in Figure 2.

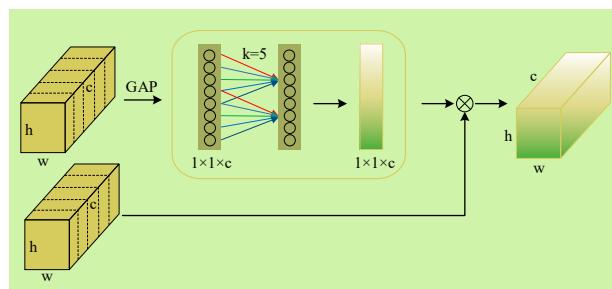


Figure 2. ECA channel attention module.

To extract the feature map $F[h, w, c]$ from the backbone network (“ h ” represents the height of the feature map, “ w ” represents the width of the feature map, and c represents the number of channels), a Global Average Pooling (GAP) operation is applied. This operation computes the average value across each channel of the feature map, resulting in a new feature map with a size of $[1 \times 1 \times c]$.

Next, a one-dimensional convolution is performed on this feature map using a convolution kernel of size 5. This convolution operation aims to obtain the weight size for each channel. The resulting weights are then multiplied by the initial feature map. The specific structure of the Global Average Pooling (GAP) operation is illustrated in Figure 3.

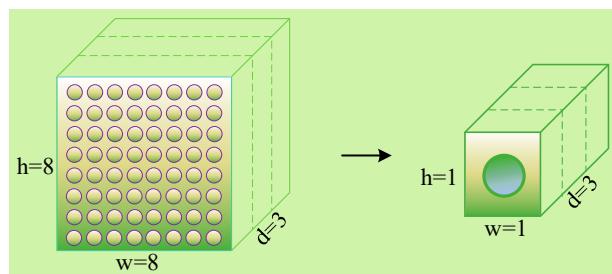


Figure 3. GAP global average pooling.

As depicted in Figure 3, for each of the three channel feature maps, there are 64-pixel points. The Global Average Pooling (GAP) operation sums up the corresponding pixel values of each point and then calculates their average to obtain three 1×1 channel maps. These maps represent the averaged values for each channel after pooling.

2.3. PConv-Based Implementation of YOLO v7 Network Model for Fast Detection

It is well known that the computation of the YOLO v7 network model consumes a lot of computational resources and, therefore, has low performance. To reduce the computational intensity and improve performance, this paper introduces a new convolution operation called the PConv convolution [28]. This operation exploits the similarity between the features of different channels of the insulator and the redundancy in its feature mapping to optimize the computational cost and reduce the number of floating-point operations. The specific structure of the PConv convolution is illustrated in Figure 4.

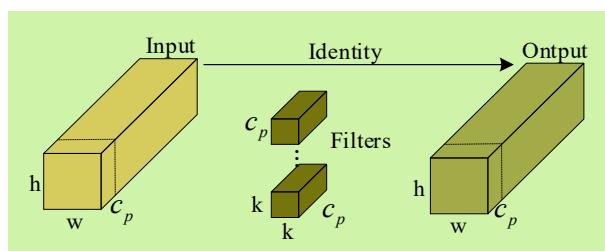


Figure 4. PConv structure diagram.

Compared to the traditional convolution operation where all channels are filtered, PConv selectively performs feature extraction only on certain input channels while leaving the remaining channels unchanged. In the case of constant insulator feature extraction in the YOLO v7 backbone network, the channels are computed as representatives of the entire feature mapping. Without loss of generality, we can assume that input and output feature mappings have an equal number of channels. The number of floating-point operations for PConv can be calculated as follows:

$$h \times w \times k^2 \times c_p^2 \quad (1)$$

where h represents the height of the feature map, w denotes the width of the feature map, k represents the convolution kernel size, and c_p denotes the number of output channels. When partially compared $r = c_p/c = 1/4$, the FLOPs of PConv are only 1/16th of those of a normal Conv in addition to the smaller memory accesses of PConv, which are only 1/4th of those of a normal Conv:

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \quad (2)$$

In PConv convolution, only c_p channels are used for spatial feature extraction, while the remaining $c - c_p$ channels stay unchanged for inter-channel feature information transfer.

In summary, by integrating PConv convolution with the traditional convolution in the YOLO v7 network model, we could enhance the detection accuracy while reducing the number of parameters in the model. This combined approach enables the effective detection of insulator faults, leading to improved overall performance.

2.4. Small Target Insulator Detection Based on NWD

There are some challenges in the dataset, such as small and dense insulators at the far end of the image, insulators at the edges of the image, and small areas of defective insulators. These situations can lead to a loss of important insulator features during the pre-feature extraction process, resulting in a lower detection accuracy for insulators. To tackle this issue, this paper proposes the use of NWD [29] as a replacement for the intersection over union (IoU) metric to improve the detection accuracy of small insulator targets.

The sensitivity of IoU varies widely for different scales of insulators, as illustrated in Figure 5. In the case of a small target measuring 25×50 pixels, box A represents the true bounding box, while boxes B and C represent bounding boxes with diagonal deviations represented by 1 pixel and 5 pixels, respectively. The IoU values for these cases are 0.88 and 0.56, indicating some degree of IoU degradation. This degradation can affect the label assignment process and subsequently impact detection accuracy.

$$IoU_{25 \times 50} = \frac{|A \cap B|}{|A \cup B|} = 0.88 \quad (3)$$

$$IoU_{25 \times 50} = \frac{|A \cap B|}{|A \cup B|} = 0.56 \quad (4)$$

In the case of normal targets with dimensions of 827×163 pixels, when the diagonal pixel deviations are the same, the IoU cross-merge ratios are 0.985 and 0.916, respectively, showing only slight variations in deviation. However, it is important to note that IoU is highly sensitive to small target position deviations. Despite YOLO v7's dynamic assignment strategy, IoU remains significantly sensitive to small target insulators. Finding a single threshold that is suitable for all cases becomes challenging, making the detection and segmentation of small target insulators more difficult.

$$IoU_{827 \times 163} = \frac{|A \cap B|}{|A \cup B|} = 0.985 \quad (5)$$

$$IoU_{827 \times 163} = \frac{|A \cap B|}{|A \cup B|} = 0.916 \quad (6)$$

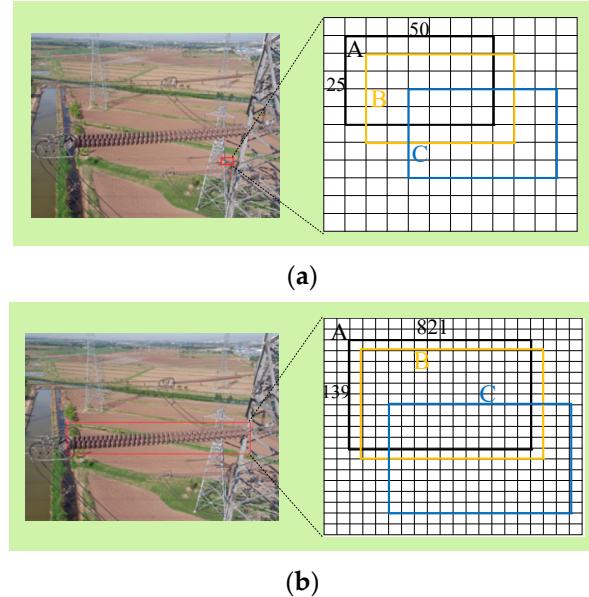


Figure 5. *IoU* sensitivity analysis graph. (a) Small target insulators. (b) Normal insulators.

To mitigate the sensitivity of the *IoU* metric to the positional deviation of small insulator targets, this paper introduces a method for detecting small target insulators based on the NWD. In the bounding box of a small target insulator, background pixels are often present, while the pixels of the detected object predominantly concentrate in the center of the bounding box. The weight assigned to each pixel gradually decreases from the center toward the boundary, with the pixel at the center having the highest weight.

To enhance the representation of pixel weights within the bounding box, this article models the bounding box as a two-dimensional Gaussian distribution.

The horizontal bounding box R is represented as (cx, cy, w, h) , where cx and cy denote the coordinates of the center point, and w and h denote the width and height of the bounding box. Based on pixel distribution characteristics, it can be expressed using the equation of an ellipse.

$$\frac{(x - u_x)^2}{\sigma_x^2} + \frac{(y - u_y)^2}{\sigma_y^2} = 1 \quad (7)$$

where u_x and u_y denote the coordinates of the center of the ellipse, $u_x = cx$, $u_y = cy$, σ_x denote the length of the x-axis, whose length is $\sigma_x = w/2$, σ_y respectively, and the length of the y-axis, whose length is $\sigma_y = h/2$.

The probability density function of a two-dimensional Gaussian distribution is provided by the following equation:

$$f(x|u, \Sigma) = \frac{\exp\left(-\frac{1}{2}(x - u)^T \Sigma^{-1}(x - u)\right)}{2\pi|\Sigma|^{\frac{1}{2}}} \quad (8)$$

where x denotes the coordinates (x, y) , u denotes the Gaussian distribution mean vector, and Σ denotes the covariance matrix. When the ellipse in Equation (7) represents the prob-

ability density function of a two-dimensional Gaussian distribution, the bounding box $R = (cx, cy, w, h)$ can be effectively modeled as a two-dimensional Gaussian distribution.

$$u = \begin{bmatrix} cx \\ cy \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \frac{\omega^2}{4} & 0 \\ 0 & \frac{\omega^2}{4} \end{bmatrix} \quad (9)$$

Next, the distribution distance is computed using the Wasserstein distance. For two two-dimensional Gaussian distributions $u_2 = N(m_2, \Sigma_2)$, the second-order Wasserstein distance is defined as follows:

$$W_2^2(u_1, u_2) = Tr \left(\sum_1 + \sum_2 - 2 \left(\sum_2^{1/2} \sum_1^{1/2} \right)^{1/2} \right) + \|m_1 - m_2\|_2^2 \quad (10)$$

The above equation is further simplified as:

$$W_2^2 = \|m_1 - m_2\|_2^2 + \left\| \sum_1^{1/2} - \sum_2^{1/2} \right\|_F^2 \quad (11)$$

The Gaussian distributions N_a and N_b for bounding box $A = (cx_a, cy_a, \omega_a, h_a)$ and bounding box $B = (cx_b, cy_b, \omega_b, h_b)$ can be expressed according to the above equation as:

$$W_2^2(N_a, N_b) = \left\| \left(\left[cx_a, cy_a, \frac{\omega_a}{2}, \frac{\omega_b}{2} \right]^T, \left[cx_b, cy_b, \frac{\omega_b}{2}, \frac{\omega_b}{2} \right]^T \right) \right\|_2^2 \quad (12)$$

where $W_2^2(N_a, N_b)$ is the W_2^2 distance metric, which cannot be used directly to detect the similarity metric between the two; therefore, the exponential form is used to normalize the new metric, called *NWD*:

$$NWD(N_a, N_b) = \exp \left(- \frac{\sqrt{W_2^2(N_a, N_b)}}{C} \right) \quad (13)$$

where C is a constant closely related to the dataset and can be set according to empirical values; in this paper, $C = 2$.

In summary, insulator detection using NMD instead of *IoU* can reduce the sensitivity of position deviation for small targets in position and improve the detection accuracy of insulators.

3. Experimental Results and Analysis

In order to comprehensively evaluate the performance of the enhanced YOLO v7 algorithm, several experiments were conducted, including attention experiments, different convolution experiments, ablation experiments, and comparisons with classical network models. The details of these experiments are as follows:

3.1. Experimental Platform

The experimental environment for the images in this paper is shown in Table 1. In this paper, the image experiment environment is Windows system, the CPU of the experiment platform is Intel i5-12400F, the graphics card is NVIDIA GeForce RTX3060, the memory is 16 GB, the input image size is 640×640 px, the batch size is 8, and the epochs are 150.

The YOLO v7 network model was trained using the PyTorch framework. This paper used a pre-trained weight model and performed model training updates on top of it. The model was trained for 150 epochs to generate the final network weight model.

Table 1. This paper experiments with hardware and software configuration.

Software	Details
OS	Windows 10
CPU	Intel i5-12400F
GPU	NVIDIA GeForce RTX3060
Memory size	16 GB
Input	640 × 640 px

The table below presents the hyperparameter settings to train the improved YOLO v7 network model, as proposed in this paper (Table 2). Momentum action reduces oscillations and noise, which is common in traditional gradient descent algorithms. Learning rates affect the speed of convergence. Weight decay reduces the risk of overfitting in the model.

Table 2. Model hyperparameter settings.

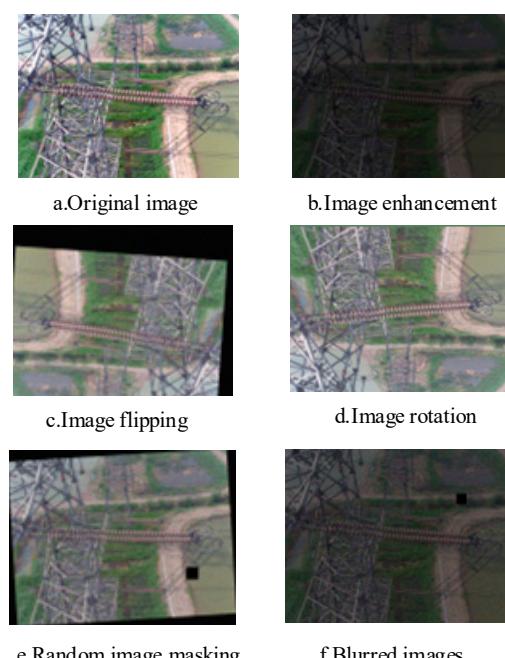
Momentum	0.937	Weight decay	0.0005
Learning rates	0.01	Batch size	8

3.2. Data Description

To evaluate the generalizability of the algorithm, experiments were conducted using two different datasets: the TISLTR dataset for normal insulators and the FISLTR dataset for high-resolution small target fault insulators.

The TISLTR dataset consists of 976 normal insulator images with varying sizes and backgrounds. The resolution of these images is 1152 × 864 pixels. On the other hand, the FISLTR dataset contains 1231 faulty insulator images with a resolution of 3216 × 2136 pixels. The faulty insulators in this dataset exhibit flash and broken fault types. The data are labeled using the Labeling tool and divided into three main categories: normal insulators, flashover insulators, and broken insulators.

The original dataset is not sufficient to satisfy the experimental generalization and universality; therefore, the dataset is expanded by color random transformation, cropping, and blurring, and the result of image data enhancement and expansion is shown in Figure 6.

**Figure 6.** Insulator image data enhancement.

To improve the generalization capability of the network model, the TISLTR and FISLTR datasets were divided into training, validation, and test sets in a ratio of 8:1:1.

3.3. Evaluation Indicators

To evaluate the superiority of the improved YO-LO v7 algorithm in an objective and accurate manner, metrics like precision, recall, and mean average precision (mAP) are employed to distinguish the network model. Precision refers to the percentage of correctly predicted positive samples using the model out of all samples predicted as positive categories. Recall is the number of samples correctly predicted by the model in the positive category as a proportion of the number of all samples actually in the positive category. The mAP is an evaluation metric in the target detection task that calculates the average accuracy value across multiple categories. Additionally, the frame rate (Frames Per Second, FPS) indicates the number of images that the model can process per second. The higher the model's processing frame rate, the quicker it can detect the images.

3.4. Experimental Results and Analysis

3.4.1. Experimental Comparison of Different Attentional Mechanisms

This study aims to verify the effectiveness of the ECA mechanism in detecting insulator targets. To achieve this objective, two other attention mechanisms, namely the Squeeze-and-Excitation networks (SE) and Convolutional Block Attention Module (CBAM), were introduced and experimentally compared with the ECA mechanism within the YOLO v7 algorithm.

The SE mechanism learns the significance of each channel's feature map through self-learning and assigns distinctive weights according to the specific characteristics. CBAM, in addition to channel attention, incorporates spatial attention by assigning varying weights to different objects or background information corresponding to feature channels and spatial locations. However, CBAM is less sensitive to smaller feature maps and is, therefore, less suitable for detecting small target faults in insulators.

Both SE and CBAM have certain limitations. On the other hand, ECA, which is built upon the SE mechanism, achieves better extraction of useful insulator features by applying one-dimensional convolution to information interaction across channels. The performance metrics of these three attention modules trained on both the TISLTR dataset (normal insulator dataset) and the FISLTR dataset (high-resolution small target faulty insulator dataset) are presented in Table 3.

Table 3. Comparison of results obtained by different attention mechanisms.

Dataset	Labels	Methods	Precision	Recall	mAP@0.5
TISLTR	insulator	SE	93.5%	90.6%	91.7%
		CBAM	96.7%	93.8%	92.1%
		ECA	90.9%	93.7%	94%
FILTER	flash	SE	74.1%	76%	80.1%
	broken		83.8%	67.9%	74.7%
	flash	CBAM	81.3%	83.4%	87.4%
	broken		85.2%	77.2%	81.3%
	flash	ECA	83.2%	89.3%	91.3%
	broken		90.2%	86.1%	88.5%

The table demonstrates that the YOLO v7 algorithm with the ECA attention mechanism outperforms both the SE and CBAM mechanisms in terms of detection results. This superiority can be observed not only in the normal insulator dataset (TISLTR) but also in the high-resolution small target fault dataset (FISLTR). These findings indicate that the ECA attention mechanism exhibits strong feature extraction capabilities for insulators within the dataset used in this study.

3.4.2. Comparison of Different Convolution Experiments

To evaluate the performance of partial convolution (PConv) in insulator target detection, this study compared it with conventional convolution (Conv) within the YOLO v7 network model. PConv achieves feature extraction for insulators by leveraging the similarity between channel features and extracting features from only a subset of the channels.

The TISLTR dataset (normal insulator dataset) and the FISLTR dataset (high-resolution small target fault insulator dataset) have been employed to evaluate the performance of both Conv and PConv. Table 4 illustrates that PConv accomplishes a higher detection accuracy while having a reduced parameter count of 4.3 MB in both datasets compared to Conv. The utilization of PConv not only decreases the count of parameters but also significantly improves the detection speed.

Table 4. Comparison of the results from various convolution experiments.

Dataset	Labels	Convolutions	mAP@0.5	mAP@0.5:0.95	FPS	Params
TISLTR	insulator	Conv	92.1%	63.4%	95.24	37.1 MB
		PConv	92.3%	66%	75.75	32.8 MB
	flash	Conv	85.9%	41.7%	131.58	37.2 MB
		PConv	86.6%	42.7%	129.87	32.8 MB
	broken	Conv	81.9%	40.6%	131.58	37.2 MB
		PConv	83.6%	45%	129.87	32.8 MB

3.4.3. A Comparison of Ablation Experiments

In order to assess the accuracy and effectiveness of the algorithms proposed in this paper and examine the impact of each module on the model's performance metrics, we conducted ablation experiments on both the TISLTR and FISLTR datasets. We used the YOLO v7 model as the base model and gradually added different modules to evaluate the performance metrics of the network model. Table 5 presents the results of the ablation experiments.

Table 5. Comparison results of ablation experiments.

Dataset	Labels	Methods	Precision	Recall	mAP@0.5	mAP@0.5:0.95
TISLTR	insultor	YOLOv7	90.9%	93.7%	92.1%	63.4%
		YOLOv7+ECA	90.9%	93.7%	94%	65.6%
		YOLOv7+ECA+PConv	91.2%	96.8%	96.7%	66.5%
		YOLOv7+ECA+Pconv+NWD	96.8%	93.7%	98.1%	68%
FITER	flash	YOLOv7	81.7%	80.4%	85.9%	41.7%
			92.6%	73.2%	81.9%	40.6%
	broken	YOLOv7+ECA	83.6%	90.2%	91.3%	47%
			89.9%	89%	90.8%	52.5%
	flash	YOLOv7+ECA+PConv	81.5%	91.6%	91.1%	48.1%
			91%	87.5%	92.5%	54.6%
	broken	YOLOv7+ECA+Pconv+NWD	87.1%	92.3%	93%	47.6%
			95.2%	90.3%	92.2%	55.2%

The experiments showed that incorporating the ECA attention mechanism into the YOLO v7 model led to an mAP enhancement of 1.9% in the TISLTR dataset and 5.4% in the FISLTR dataset. This suggests that the ECA attention mechanism allows the network model to concentrate on vital channels and have particular extraction capabilities to detect insulators in complex background images.

When comparing the performance metrics of YOLO v7+ECA and YOLO v7+ECA+PConv, the results indicate that the inclusion of PConv enhances Precision and Recall slightly, while

significantly improving the average detection accuracy. These findings suggest that PConv improves the network model's ability to extract features of insulators by utilizing channel similarity in conjunction with the channel attention provided by ECA.

In conclusion, the proposed YOLO v7+ECA+PConv+NWD network structure combines the strengths of the ECA attention mechanism, partial convolution (PConv), and normalized Wasserstein distance (NWD) in insulator target detection. This combination significantly enhances the performance metrics of the single-stage network model YOLO v7 in power system insulator detection, including higher Precision, Recall, and mAP scores.

To further demonstrate the ability of the improved network model to extract insulator features, heat maps of the model were generated using both the TISLTR and FISLTR datasets, which were trained with the best.pt model. As shown in Figure 7, the heat maps demonstrate that the improved network model effectively highlights relevant features of insulators. This capability allows for the detection of insulator defects within power systems.

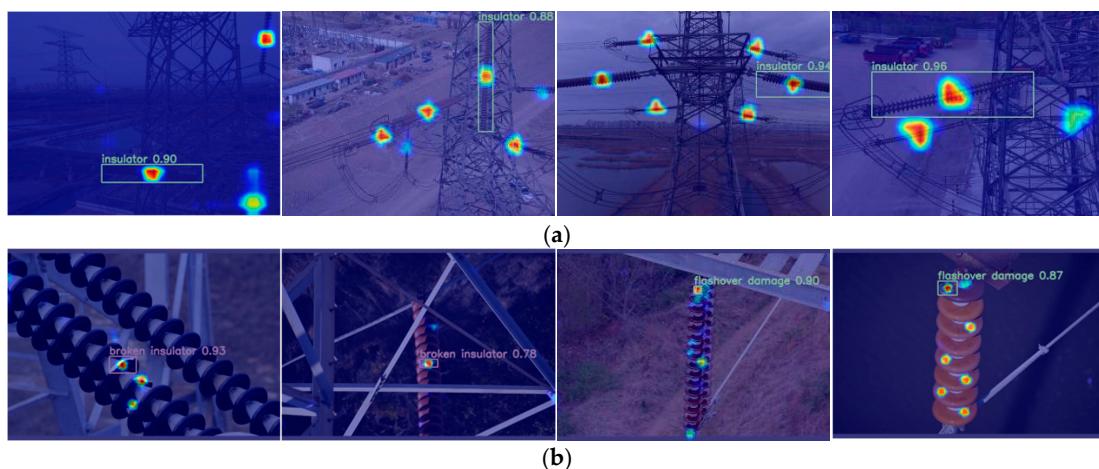


Figure 7. This paper improves the YOLO v7 network model insulator heat map. (a) Normal insulator thermal diagram. (b) Thermal diagram of broken, flashing insulators.

3.4.4. Experimental Comparison of Different Network Models

The improved YOLO v7 algorithm proposed in this paper was evaluated using the TISLTR dataset, and its performance was compared against several currently mainstream target detection network models, such as the single-stage network models SSD, YOLO v3, YOLO v4, YOLO v5, YOLOx, YOLO v7-tiny and the two-stage network model Faster R-CNN. Table 6 displays the relevant performance metrics obtained from the experiments.

Table 6. Detection results of different network models.

Methods	Precision	Recall	mAP@0.5
SSD	92.36%	59.92%	78.18%
YOLOv3	93.47%	76.86%	88.31%
YOLOv4	92.00%	57.02%	75.69%
YOLOv5	91.7%	87.5%	89.3%
YOLOx	88.56%	86.36%	90.43%
FasterR-CNN	53.92%	90.91%	86.9%
YOLOv7-tiny	91.82%	83.47%	92.01%
Ours	96.8%	93.7%	98.1%

As seen in Table 6, both the traditional single-stage network models and the two-stage network model performed well in power system insulator target detection. The mAP (Mean Average Precision) of all network models exceeded 75%, with the improved YOLO v7 network model in this paper achieving an mAP of 96.2%, which is higher than that of

other network models. However, SSD and YOLO v4 had the lowest Recall, as indicated by Figure 8, which shows instances of missed detections and leakage detection. In contrast, the improved network model in this paper achieved a Recall of 93.7% and a Precision of 96.8%, both of which surpass other network models. YOLO v7-tiny, being a lighter version of YOLO v7, demonstrated a detection performance second only to that of the improved YOLO v7 network model in this paper. The detection results of YOLO v3, YOLO v5, YOLO x, and Faster-CNN show that their network models are all weaker than the improved network model in this paper. Therefore, the improved network model proposed in this article exhibits a certain superiority.

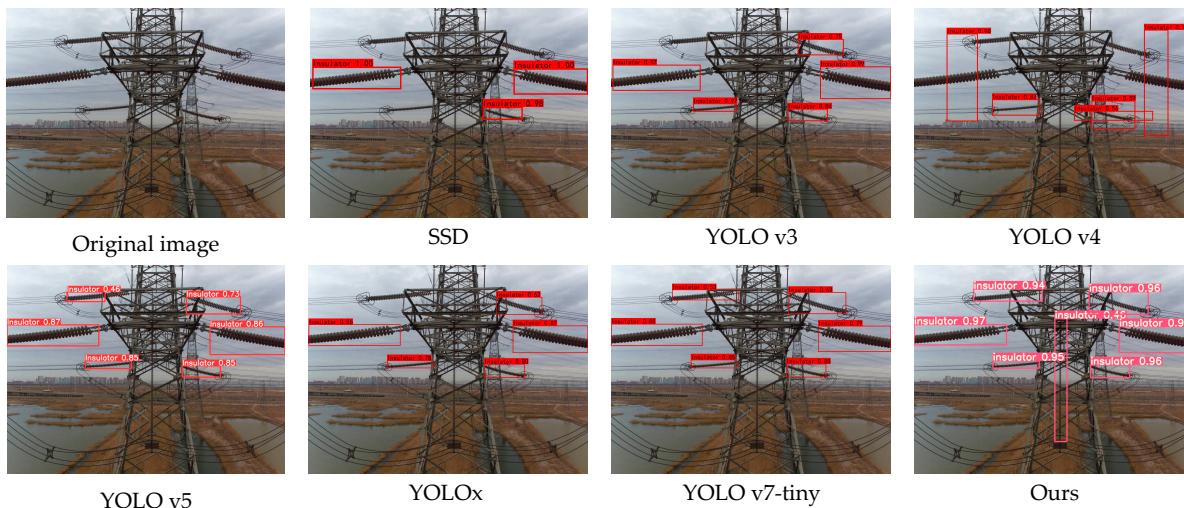


Figure 8. Comparison of insulator detection results of different network models in a complex background.

Figure 8 illustrates that insulator detection is challenging due to the presence of background utility poles. The middle vertical insulator, in particular, almost blends into the background of utility poles, resulting in a complex background for insulator detection. Additionally, Figure 9 presents the detection results for different insulator sizes. It can be observed that other algorithms missed detections during the process, while the improved network model from the article successfully detected insulators in complex backgrounds. The improved network model demonstrated better results for both insulator detection in complex backgrounds and the detection of insulators with different sizes.

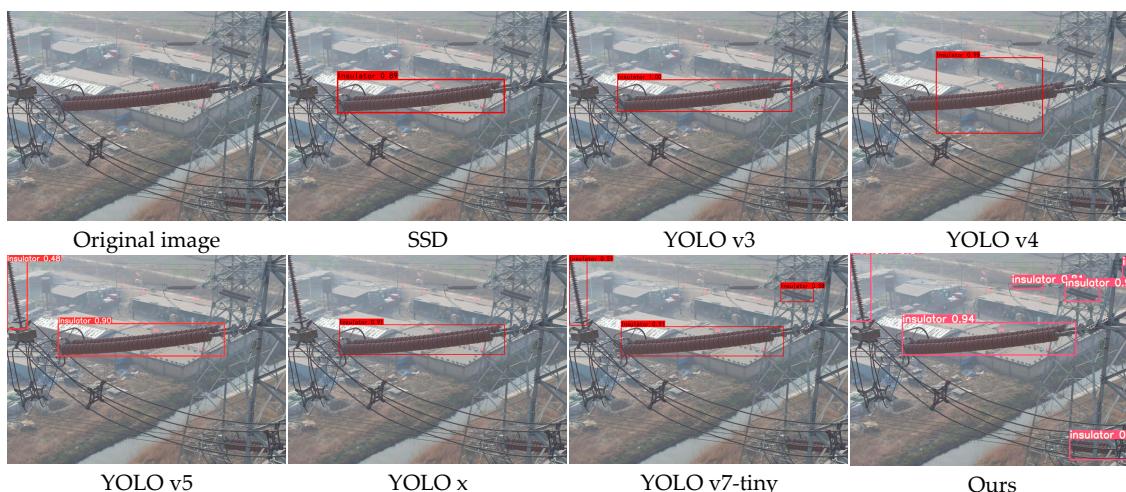


Figure 9. Comparison of insulator test results of different sizes.

4. Conclusions

- (1) This paper proposes a model based on the YOLO v7-ECA-PConv-NWD network for identifying normal insulators, broken insulators, and flashover insulators in power transmission lines. After training, the model achieved detection accuracies of 98.1%, 93%, and 92.2%, respectively, resulting in enhanced performance. This provides a theoretical basis for the intelligent maintenance of power systems.
- (2) This paper proposes adding an ECA mechanism to the backbone of the YOLO v7 network model to improve its attention toward insulator features and reduce noise interference. As a result of this improvement, the accuracy of insulator detection could be enhanced.
- (3) A new PConv convolution was utilized in this paper to reduce the number of floating-point operations while processing insulators in the target detection task. By exploiting the similarity between different channel features and the redundancy in feature mapping, the PConv convolution successfully reduced the parameter size by 4.3 MB compared to conventional convolution. Despite this reduction, the accuracy of insulator detection remained largely unaffected.
- (4) This paper utilized NWD in place of the IoU ratio for target detection. By substituting NWD for IoU , the model became less susceptible to position deviations in small target insulators. Based on the fact that the center pixel has the highest weight and decreases from the center to the edge, it addresses the challenge of high miss detection rates in images.

This paper only achieved the detection of insulator breakage and flashover of transmission lines, and it is relatively limited in terms of fault types. The next stage is to broaden the scope of fault types and incorporate the enhanced network model into DWF3.0, a big data application development tool developed by the Tsinghua National Laboratory of Big Data System Software. By doing so, we aim to construct a more comprehensive system for detecting and providing alerts in insulator failures.

Author Contributions: Methodology, J.Z., X.W., L.Z., L.Y., Y.C. and M.T., conceived and designed the experiments, J.Z. and X.W., performed the experiments, J.Z. and L.Z.; conceptualization and investigation, M.T. and L.Y.; analyzed the data, X.W. and L.Z.; wrote the paper, J.Z. and X.W.; funding acquisition, L.Z. and Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Xinjiang Uygur Autonomous Region, grant number 2022D01C431, and Xinjiang Uygur Autonomous Region Youth Science Foundation Project, grant number 2022D01C693.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author.

Acknowledgments: The authors want to thank the editor and anonymous reviewers for their valuable suggestions for improving this paper.

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

References

1. Mao, J.; He, Z.; Wang, Y.; Zhang, H.; Zhong, H.; Yi, J.F.; Tao, Z.M.; Chen, N.T. A review of research and applications on path planning technology for power inspection robots. *Control. Decis.* **2023**, *1*–16. [[CrossRef](#)]
2. Liu, C.; Wu, Y.; Liu, J.; Han, J. MTI-YOLO: A lightweight and real-time deep neural network for insulator detection in complex aerial images. *Energies* **2021**, *14*, 1426. [[CrossRef](#)]
3. Zhao, L.; Wang, P.; Yao, H.; Tian, M. Survey of power line extraction methods based on visible light aerial image. *Power Syst. Technol.* **2021**, *45*, 1536–1546.
4. Yuan, W.H.; Liu, B.; Wang, Z.; Sai, T.; Xia, X.; Wang, F.; Zhong, L. Denoising of insulator hydrophobic image based on the guided filtering algorithm. *J. Electr. Power Sci. Technol.* **2021**, *36*, 135–140.
5. Li, H.R.; Li, Z.H.; Wu, T.; Xu, X.R.; Fang, M.Q. Research on edge detection of insulators by canny operator optimized by two-dimensional maximum entropy threshold. *High Volt. Appar.* **2022**, *58*, 205–211+220.
6. Huang, X.; Liu, X.; Zhang, Y.; Li, J.; Zhang, H.; Xing, X. Classification recognition method of the insulator in aerial image based on the red-blue difference and developed K-means algorithm. *High Volt. Eng.* **2018**, *44*, 1528–1534.

7. Chen, G.C.; Gu, G.M.; Yu, X.N.; Li, Z.-b. Method for defect detection of catenary insulators by combining the features of PHOG and BOW-SURF. *J. Chin. Comput. Syst.* **2021**, *42*, 172–177.
8. Mei, X.; Chen, J.; Wang, S.Y. Composite insulator hydrophobicity image detection method based on morphology. *High Volt. Eng.* **2013**, *39*, 3088–3094.
9. Wu, Y. Research on Insulator Recognition Methods in Aerial Images Based on MACHINE Learning. Ph.D. Thesis, North China Electric Power University, Beijing, China, 2016.
10. Huang, H.H. The Research and Implementation of Anomaly Detection System on High-Speed Rail Catenary Insulators Based on Multi-Core Parallel Technology. Ph.D. Thesis, Xi'an University of Electronic Technology, Xi'an, China, 2018.
11. Liu, K.P.; Li, B.Q.; Qin, L.; Li, Q.; Zhao, F.; Wang, Q.L.; Xu, Z.P.; Yu, J.Y. Review of application research of deep learning object detection algorithms in insulator defect detection of overhead transmission lines. *High Volt. Eng.* **2023**, 1–12.
12. Gkioxari, G.; Hariharan, B.; Girshick, R.; Malik, J. R-cnns for pose estimation and action detection. *arXiv* **2014**, arXiv:1406.5212.
13. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)]
14. Zheng, R.J. Research and application of damaged insulator recognition based on R-CNN. Ph.D. Thesis, Hubei Minzu University, Enshi City, China, 2021.
15. Yi, J.Y.; Chen, C.F.; Gong, G.Q. Aerial insulator detection of transmission line based on improved Faster RCNN. *Comput. Eng.* **2021**, *47*, 292–298+304.
16. Tian, Z.J.; Wu, J.Q.; Zhang, W.Q.; Chen, W.; Yang, W.; Wang, S. An insulator faults detection method in low illuminance complex environment based TOE-Net. *Power Syst. Technol.* **2023**, 1–14.
17. 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.
18. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.
19. Redmon, J.; Farhadi, A. YOLOv3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.
20. Zan, W.; Dong, C.; Zhao, J.; Hao, F.; Lei, D.; Zhang, Z. Defect Identification of Power Line Insulator Based on an Improved yolov4-tiny Algorithm. In Proceedings of the 5th International Conference on Renewable Energy and Power Engineering (REPE), Beijing, China, 28–30 September 2022; pp. 35–39.
21. Zhu, X.; Liu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 2778–2788.
22. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. SSD: Single Shot Multibox Detector. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Part I 14, Amsterdam, The Netherlands, 11–14 October 2016; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 21–37.
23. Zhu, Y.C.; Zheng, Y.; Qin, J.L. Insulator target detection based on improved YOLOv3. *Insul. Surge Arresters* **2022**, *2022*, 166–171.
24. Song, L.B.; Fei, Y.Q. A New YOLOv4-tiny neural network and its application on object detection of power-line insulators. *J. Harbin Univ. Sci. Technol.* **2022**, *27*, 73–79.
25. Wang, J.Y.; Xu, X.Y.; Xie, G.; Yan, G.W. Insulators self-explosion fault detection based on improved SSD model. *Mod. Electron. Tech.* **2022**, *45*, 115–121.
26. Wang, C.Y.; Bochkovskiy, A.; Liao HY, 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; pp. 7464–7475.
27. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11534–11542.
28. Chen, J.; Kao, S.; He, H.; Zhuo, W.; Wen, S.; Lee, C.H.; Gary Chan, S.H. Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 12021–12031.
29. Wang, J.; Xu, C.; Yang, W.; Yu, L. A normalized Gaussian Wasserstein distance for tiny object detection. *arXiv* **2021**, arXiv:2110.13389.

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.