



Insulator Defects Detection for Aerial Photography of the Power Grid Using You Only Look Once Algorithm

Shanshan Wang¹ · Xinyi Zou¹ · Wei Zhu¹ · Liang Zeng¹

Received: 11 July 2022 / Revised: 11 December 2022 / Accepted: 5 January 2023 / Published online: 18 January 2023
© The Author(s) under exclusive licence to The Korean Institute of Electrical Engineers 2023

Abstract

The safe operation of the power grid system depends partly on regular inspections of transmission lines, in which the insulator is one of the most important inspections objects. The manual inspection of transmission lines is a chaotic process that is both time and cost-consuming since it involves an accurate manual inspection by an expert. For insulator defect detection, an improved YOLOv4 algorithm is proposed. First, a new data augmentation method is proposed to solve the problem of insufficient sample size. Then, the size of the anchor boxes is redesigned base on the K-means algorithm to further improve the detection precision. Finally, an insulator defects detection network is constructed based on YOLOv4. Experimental results show that the detection precision of the improved network is 37.2% higher after data enhancement and anchor boxes redesign. In addition, the detection method proposed in this paper is superior to other popularity detection algorithms, including the single shot detector, region-convolutional neural networks (Faster-RCNN) and released version of you only look once (YOLO). The value of mean average precision is 99.08% and frame per second is 56. The robustness test results demonstrate that our proposed algorithm performs well under different light intensities and complex environmental backgrounds, and can accurately detect all targets, which is significantly better than other comparative algorithms. In terms of detection accuracy, test speed and robustness, our proposed algorithm meets the requirements of industrial field applications of insulator defect detection.

Keywords Insulator defects detection · YOLO algorithm · Data augmentation · Object detection · Aerial images

1 Introduction

The safe and stable operation of transmission lines is the most basic demand of society, and insulators are an important part of transmission lines. Insulators in outdoor environment for a long time are prone to breakage, self-explosion, string falling and other defects. These insulator defects will lead to a series of serious problems such as the interruption of transmission lines and the collapse of power systems. Therefore, insulator defect detection is very important for the stability of the transmission system, and it is also a

challenging task [1, 2]. Therefore, automatic processing of insulator defects detection base on aerial images is a prevalent choice. The inspection methods for insulator defects mainly include manual inspection, helicopter inspection [3], robot inspection [4], and drone inspection [5]. The aerial images are taken by helicopters or unmanned aerial vehicle (UAVs), which are sent back to the computer for image processing to assist the detection of technical personnel [6]. Manual detecting of insulator defects is low efficiency, high costs, and prone to fatigue errors. In recent years, with the rapid improvement of machine vision, technologies of object detection base on machine learning have been used extensively in the fields of biological images, agriculture, and human-computer interaction [7–10].

The insulator defect detection algorithms can be divided into two categories: the traditional machine learning algorithms and the deep learning algorithms. Generally, the traditional machine learning algorithms extracted features based on image processing techniques, connected a machine learning-based detection algorithm, such as support vector

✉ Liang Zeng
zengliang@hbut.edu.cn

Xinyi Zou
243263738@qq.com

¹ School of Electrical and Engineering, Hubei Engineering Research Center for Safety Monitoring of New Energy and Power Grid Equipment, Hubei University of Technology, Wuhan, Hubei, China

machines, template matching, and adaptive thresholds. Combining the color features with spatial features, an insulator defects detection algorithmic based on spatial morphological features is proposed in [11]. A laser online monitoring method s proposed in [12] to detect insulator conditions and predict the occurrence of flashover. A fusion algorithm is proposed in [13], which combines shed contour features and ray similarity matching. A composite insulator defect detection method based on Hough transform ellipse detection is proposed in [14], the Canny operator is used to extract the edge of the insulator, and after elliptic curve fitting, it is compared with the edge curve of the real insulator to detect the contour of the insulator. The above detection methods are based on traditional machine learning algorithms and belong to the category of shallow learning architectures. It is difficult to detect defects using traditional machine learning algorithms under poor quality images, which will affect the accuracy and robustness of detection.

In the area of detection, deep learning-based object detection algorithms have been at the center of attention in numerous machine vision tasks [15, 16]. In general, deep learning algorithms are composed of a multilayer neural network that learns features directly from the highly nonlinear data without any manual feature extraction., Region-convolutional neural networks (Faster-RCNN) [17], single shot detector (SSD) [18], and you only look once (YOLO) [19] are the most popular deep learning networks used in the insulator defect detection tasks. The Faster R-Transformer model is proposed in [20], which combines a self-attention mechanism and a convolutional neural network for insulator detection.

In [21], the residual neural network (Resnet) network is combined with the Faster R-CNN network and connect to a fully convolutional network (FCN) to inspected the insulator. However, the false detection rate will increase in the above methods, due to the greatly varied in the surrounding detection environment. In [22], a detection network combining the MobileNet light network and YOLOv4 network was proposed, and the detected mAP value reaches 93.81%. In this paper, the following problems will be solved to realize the insulators defect detection: (1) The generalization capability of

algorithms should be improved to overcome the interference of the complexity of aerial images; (2) Try to improve the detection accuracy and the detection speed together. An improved method base on you only look once vision4 (YOLOv4) [23] network is proposed for insulator defect detection of the aerial images in highly diverse outdoor environments.

The remainder of this paper is arranged as follows. The proposed insulator defects detection method is elaborated in Sect. 2. The experimental details and results compared with other algorithms are shown in Sect. 3.2. In Sect. 5, we discuss the training strategy of the network model and analyze the detection effect in various scenarios. Finally, Sect. 5 is devoted to the conclusion.

2 Materials and Methods

2.1 Proposed Approaches

The China power line insulator data set (CPLID) [24] used in this paper is provided by the state grid corporation of China, and the dataset has been published online. There are 848 high-resolution images, each with a resolution of 1152×864 pixels. The data set consists of two categories, 600 ordinary insulator images and 248 defective insulator images. These images were taken by UAVs from different angles. Figure 1 shows some images from CPLID. In Fig. 1, from left to right, there are different insulators on power pylons erected on water, in forests and in villages.

However, the CPLID has contained numerous actual power transmission systems scenes, the number of defects images is insufficient for defect detection. In this section, a new data augmentation method composed of the affine transformation and the mosaic is proposed. The CPLID is expanded, the original images are set as the test set, and the data-enhanced images are set as the training set. The operation details are described below.

The dataset could be augmented directly through generally affine transformations, such as translation, scaling, and rotation. In CPLID, the location of insulators is

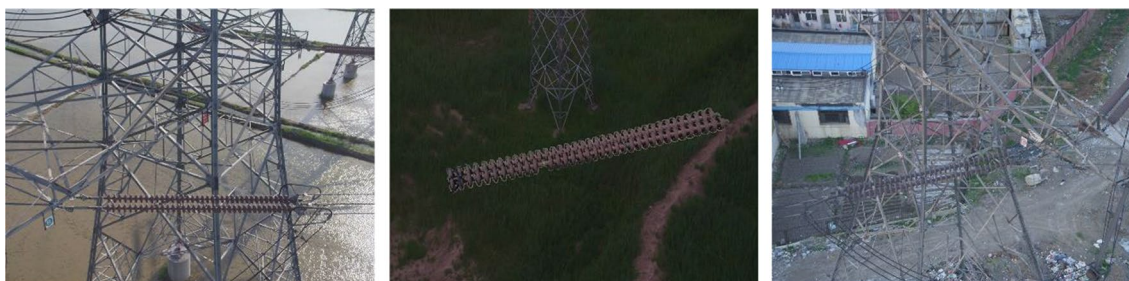


Fig. 1 Typical images in the CPLID

mainly distributed at the center of the aerial image and occupying a large area. Therefore, only rotation operation is adopted to construct a new dataset to avoid losing objects during other operations. The image rotates at a random angle around its center point. The rotation operation can be calculated as:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

where the location of a pixel for the original image and the transformed image is defined as ‘(x, y)’ and ‘(u, v)’, respectively, the θ represents the rotation angle. The image after random angle rotation is shown in Fig. 2. For example, Fig. 2a is an image with $\theta=0^\circ$, and Fig. 2b shows another image with $\theta=330^\circ$.

After that, the transformed images are processed through the mosaic method, a data augmentation method proposed by YOLOv4. That mixes four images to enrich the background of the detected object. As elaborated in Fig. 3, the four transformed images are cropped, scale, randomly spliced into one image, and reshaped to a specific size. In this Fig, the white frames indicate the objects of insulators to be detected.

There are various limitations to training the network under the default size of the anchor boxes, such as the low-performance detection and predict boxes size mismatched with the actual object. The anchor box determined by prior knowledge can significantly improve the detection performance. Thereby, the size of the anchor boxes is redesigned base on the K-means, an algorithm for clustering analysis by the iterative solution.

The size of training images has been reshaped to 416×416 . Three feature maps are extracted, which are 13×13 , 26×26 , and 52×52 , each feature map set three anchor boxes. The ratio of width to height is regarded as cluster object. All the ratio data are divided into nine groups to selecting the cluster center for each group by multiple clustering. The average of all clustering results is set as the final value of the anchor box size for the network training. The values of the anchor box size are set to (23,22), (86,21), (112,35), (186,46), (251,180), (279,87), (291,47), (293,65), and (294,121), respectively.

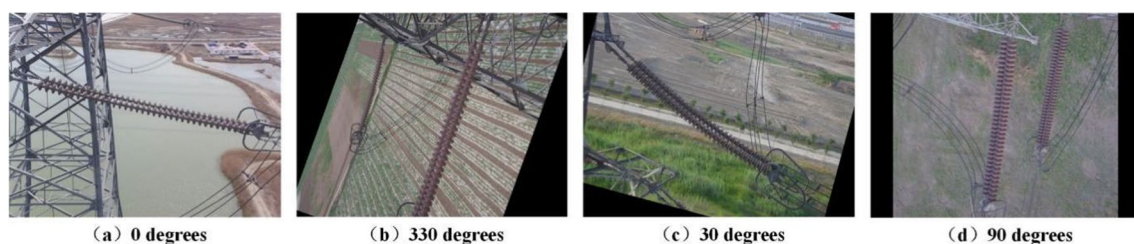


Fig. 2 Affine transformation for different images



Fig. 3 Example of images with data augmentation

2.2 Insulator Defect Detection Method

The YOLO network is an end-to-end object detection algorithm proposed in 2016 that can directly regard the detection task as the regression task. The advantage of the algorithm is that the context information of images can be extracted, which will effectively promote the accuracy and detection speed. The schematic diagram of insulator defects detection using YOLO network is shown in Fig. 4. As shown in Fig. 4a, the input insulator image is divided into a 7×7 grid, and we need to detect 9 bounding boxes for each small grid. Furthermore, each bounding box contains 5 values: x, y, w, h and confidence score. Specifically, x, y, w and h represent the coordinates of the central position of the target to be detected and its height and width, respectively, and the confidence score denotes the probability that the detected target belongs to the defect or insulator. Figure 4b draws all the priori boxes ($7 \times 7 \times 9$) of the insulator image. Some of these boxes are thicker, and some are thinner, which is an indication of different confidence levels, with thicker borders for higher confidence and thinner borders for lower confidence. Figure 4c depicts the target classes to which the different grids belong, with purple indicating the image background, yellow indicating insulator string, and red

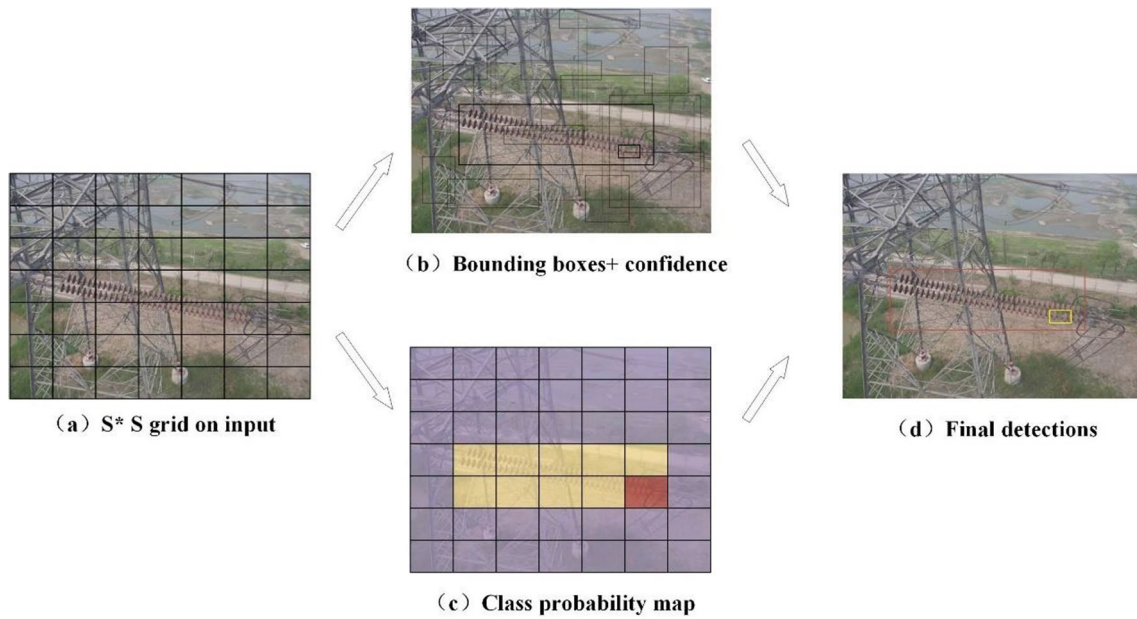


Fig. 4 Schematic diagram of insulator defects detection using YOLO network

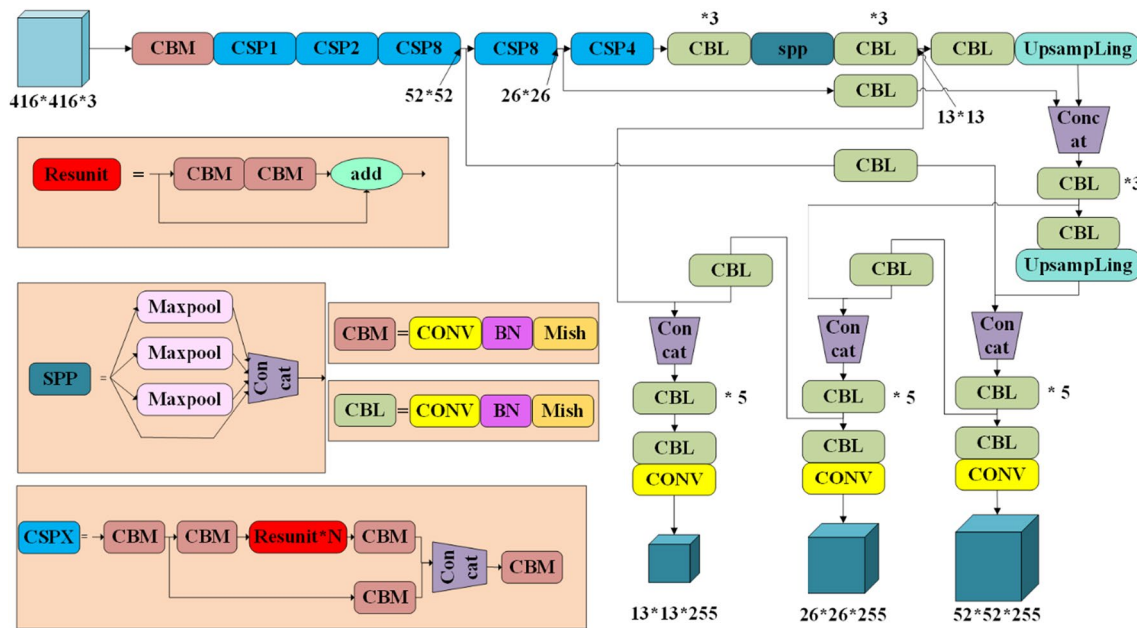


Fig. 5 Schematic diagram of YOLOv4 network

indicating insulator defect. Finally, the location of the defect target is found by the maximum suppression and confidence comparison, and it is framed in the input insulator image as shown in Fig. 4d.

The network structure of YOLOv4 is shown in Fig. 5. When YOLOv4 is used to detect insulator defects, the image is first input into YOLOv4's backbone network cross-stage-partial-connections (CSP) Darknet53, which is improved on

the basis of YOLOv3 backbone network [25]. The structure of CSPnet [26] is added to enrich the insulator image characteristics and improve the detection accuracy of the network. Secondly, the feature map of the insulator image is generated by the convolution block of multi-layer CBL, which consists of convolution (Conv) layer, batch normalization (BN) layer, and Leaky Relu layer. Thirdly, the feature map passes through the spatial pyramid pooling network

(SPP-Net), which increases the receptive field of the image [27]. Then, the top-down feature pyramid network (FPN) [28] is used to convey semantic features, and the bottom-up path aggregation network (PAN) [29] is used to convey positioning features, so as to enhance the ability of the network to extract insulator features. Finally, the prediction results of YOLOv4 are output.

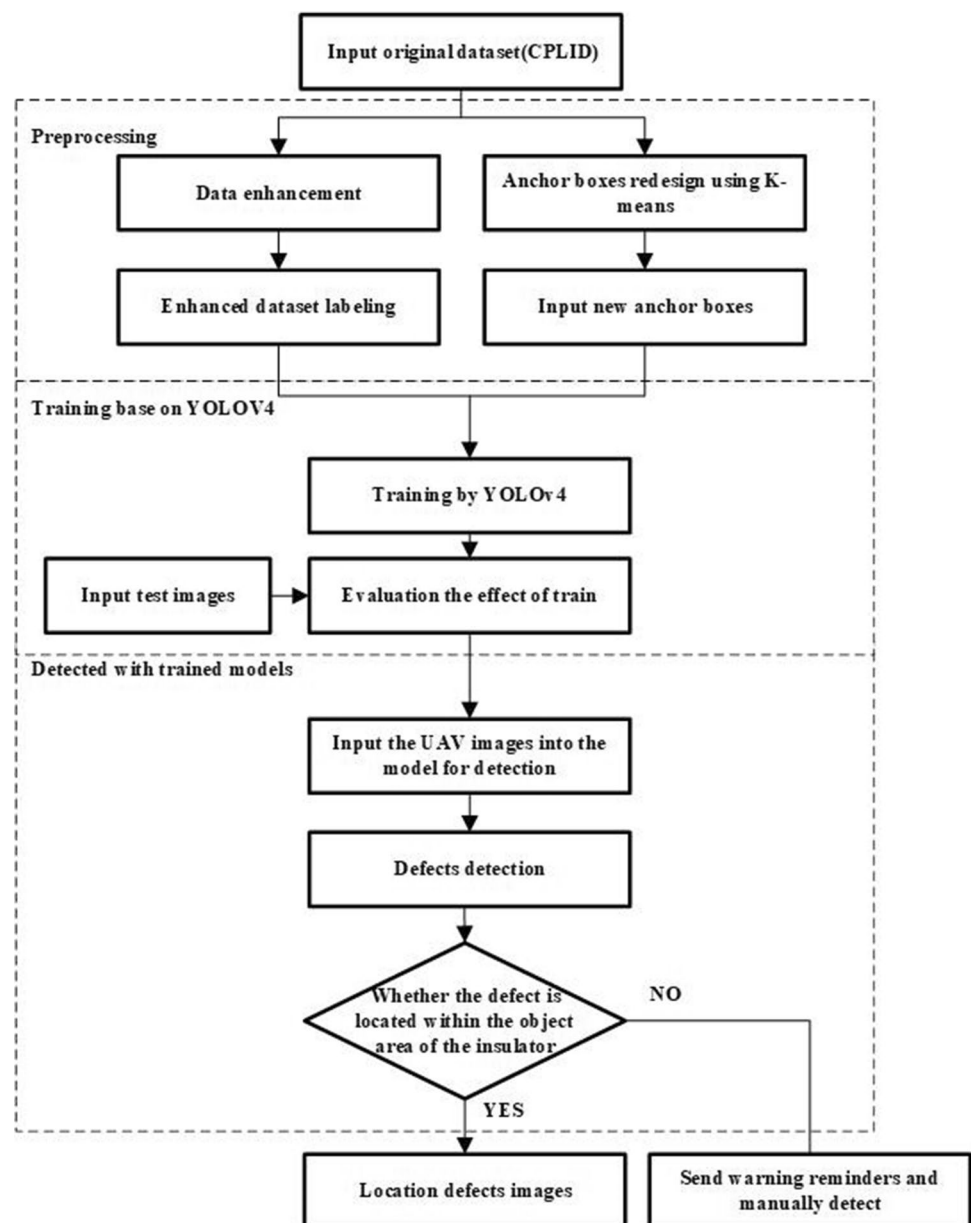
Next, this section introduces the framework of the power line insulator defect detection procedure. The flow chart is shown in Fig. 6.

First, the original dataset is processed through data enhancement, the size of anchor boxes has been redesigned

base on k-means arithmetic. Second, the training set is input into the network to update the optimal model. The effect of the trained model is evaluated by testing set. Finally, the images are input into the trained model for detection. Moreover, to verify whether the defects obtained from the model are labeling correctly in the image, the defect should be located in the area of the insulator object. If the location of the defect is outside of the insulator area, it is considered as the error-detection result. A warning message will be sent and the image will be examined manually.

The pseudo code of the feature extraction algorithm in this paper is as follows:

Fig. 6 The flowchart of the insulator defect detection



 Insulator defect detection algorithm

Input: *image dataset* the original dataset(CPLID)

Output: *images* The insulator image of the defects were located.

```

1: enhanced dataset ← Data enhancement (dataset)

2: anchor boxes size ← K-means (dataset)

3: weight file ← trained by YOLOv4 (enhanced dataset, anchor boxes size)

4: detected image ← detected by YOLOv4(weight file, UAV image)

5: for each detected image do
6:   if the defect is located within the object area of the insulator
7:     message ← detect success
8:   end if
9:   else
10:    message ← detect errors
11:  end else
12:  return detected image
13: end
  
```

2.3 Evaluation Metrics

Five common indices, precision, recall, F1 score, mean average precision (mAP), and frame per second (FPS) are used to evaluate our defect detection method. The criterion for object detection prediction is that the coincidence rate between the bounding box and the actual box is greater than the threshold, namely, the value of intersection-over-union (IoU) is greater than the threshold. The intersection-over-union (IoU) represents the coincidence rate between the predicted box and the actual box. The formulas for calculating the accuracy and recall rate are as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

where TP is the number of objects detected correctly, FP is the number of objects detected incorrectly, FN is the number of objects for leak detection. The precision and the recall are a pair of contradictory indicators, one of which increases and the other decreases. Therefore, the F1 score is used to balance the two parameters. The calculation formula of F1 is as follows:

$$F1 = \frac{2 \times P \times R}{P + R} \quad (4)$$

The AP value is the area under the precision-recall curve. Generally, models with good detection performance have higher AP value. The mAP is the average AP of all detected objects, which measures the performance of the model in all categories. The formula can be described as:

$$AP = \int_0^1 P(R)dR \quad (5)$$

$$\text{mAP} = \frac{1}{n} \sum_{i=1}^n AP_i \quad (6)$$

The FPS denotes the number of frames per second. The FPS is affected by hardware conditions, so it is generally tested under the same hardware conditions.

3 Experiment and Results

The data enhancement method is used to expand the CPLID from 848 to 2544 images, including 1800 normal insulator images and 744 defect images. The deep learning framework is TensorFlow. Other configurations of the system include the NVIDIA Quadro P2200 GPU, Inter Core i9-9900K

CPU 3.6 GHz, GPU acceleration library 10.1, and Operating system Windows 10. In the paper of YOLOv4, the accuracy is improved by introducing the features included mosaic data augmentation, cosine annealing scheduler, and label smoothing. By studying the impact of different features during training (detailed information is described in the section of the discussion), only the mosaic data enhancement is adopted in this task.

The features extracted in the backbone layer are universal. At the beginning of training, freeze the backbone network layer so that more memory will be used to train deeper layers. The training speed of the model will decrease. It is inspired by transfer learning, the training phase is divided into two-stage, namely the freezing stage and the thawing stage. During the freezing stage, the initial learning rate is set to 0.001 and the maximum learning rate is set to 0.01. During the thawing stage, the initial learning rate is 0.0001 and the maximum learning rate is 0.001. During network training, the model iterates for 25 epochs respectively, and the batch size is set to 2.

3.1 Comparison with Other Defect Detection Methods

This section verifies the performance of the model in terms of detection accuracy, calculation, and detection speed. The proposed method and other models are detected under the same data set, including SSD, Faster-RCNN, YOLOv3 [30], YOLOv4-tiny and YOLOv5× [31]. The SSD is a typical one-stage target detection algorithm, which adopts an end-to-end detection method. The Faster-RCNN is a popular two-stage object detection algorithm with high detection accuracy. The YOLOv3 is also the two-stage object detection algorithm with real-time performance. The YOLOv4-tiny is a streamlined version of the YOLOv4, which reduces the number of layers requiring feature fusion. YOLOv5 is released following YOLOv4. The YOLOv5 network includes four versions YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5×. The detection results of SSD, Faster-RCNN, YOLOv3, YOLOv4-tiny, YOLOv5× and the proposed method are shown in Table 1.

The YOLOv5× and the YOLOv4(ours) achieved high scores in mAP with 99.01% and 99.08%, respectively. The precision value of YOLOv4(ours) is 91%. Meanwhile, the AP value of the defect is 100%, which indicates that all defects have been detected. The F1 value of YOLOv4(ours) is also significantly better than other algorithms. The recall of YOLO v5×, v4 is 91.98 and 98.84, and the precision of YOLO v5×, v4 is 99.18 and 91, respectively. Simultaneously achieving high values of precision and recall is very difficult in practice, and we need to choose which one is more important for our task. Usually, an increase in precision is accompanied by a decrease in recall, and vice versa.

The Faster R-CNN network is a typical two-stage object detection network, which needs to generate candidate regions of size $W \times H \times K$ through the region proposal network (RPN) operation, thus its FPS value is low. Moreover, the ROI-Pooling layer in the Faster R-CNN network is a fully connected layer, which will generate many repeated operations when training the network, thereby reducing the training speed of the model. All the feature information used for classification and localization of SSD comes from different feature layers, and the size of the a priori frame is fixed, which causes the detection result to not match the target size in the image, and the recognition and localization of small targets are not accurate enough. Therefore, the results in Table 1 show that Faster R-CNN and SSD have lower AP values for defects.

The specific experimental results are shown in Fig. 7, the precision and recall curves of YOLOv4(ours) are shown in Fig. 8. Similar to the numerical results presented in Table 1, the superiority of YOLOv4(ours) in several metrics such as Recall, mAP@0.5 and F1 is clearly depicted in Fig. 7. Specifically, SSD and Faster-RCNN perform poorly on mAP@0.5 and F1, with the corresponding numerical results below 50, while YOLOv4(ours) approaches 100. Our algorithm behaves similarly to YOLOv5 on two contradictory metrics, Precision and Recall, but it is superior to YOLOv5 on Recall. In terms of detection speed, our algorithm and YOLOv5 achieve similar level of performance, both around 60 frames per second. The experimental results show that the detection algorithm based on YOLOv4 proposed in this paper is more suitable for

Table 1 Detection performance of several detection models

Model	Precision (%)	Recall (%)	AP		mAP @0.5 (%)	F1	FPS
			Defects (%)	Insulator (%)			
SSD	92.7	54.96	0.81	66.97	33.89	40.5	28
Faster-RCNN	61.16	76.11	2.18	86.38	43.19	41.5	25
YOLOv3	78.15	97.5	96.62	97.57	97.09	92.5	31
YOLOv4-tiny	69	97	71.18	98.78	84.98	74.5	89
YOLOv5×	99.18	91.98	99.5	98.51	99.01	95	62
YOLOv4(ours)	91	98.84	100	98.17	99.08	98	59

Bold values are the best results of the evaluation indicators

Fig. 7 The performance of various algorithms

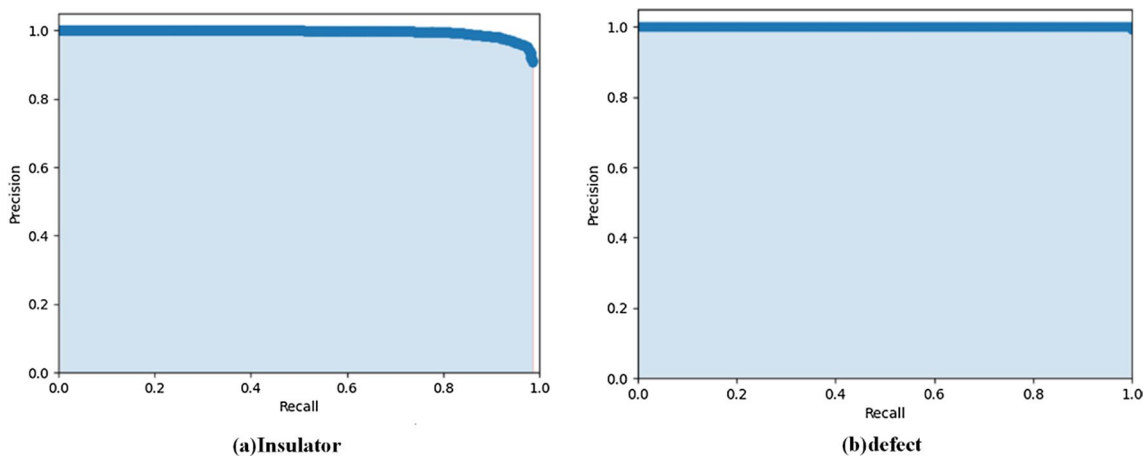
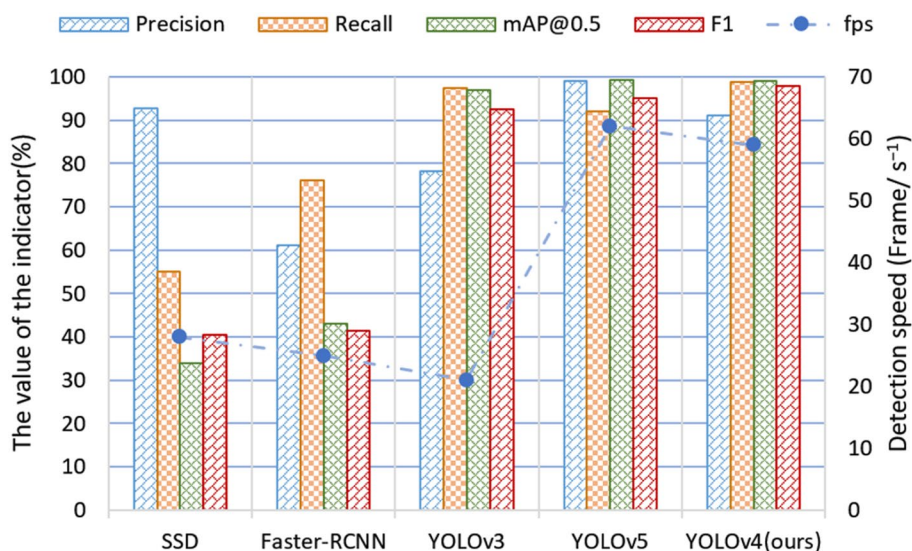


Fig. 8 The precision and recall curves of insulator defects detection

Table 2 Comparisons between our results and previous works.

No.	Company or organization	Detected network	F1	Recall (%)	mAP (%)	Precision (%)
1	Hefei institutes of physical science	YOLOv5 [15]	Not given	Not given	99.05	86.8
2	Zhengzhou University	Faster R-CNN+ Resnet-101 [21]	Not given	Not given	Not given	96.83
3	Wuhan University	MobileNet+ YOLOv4 [22]	Not given	Not given	93.81	97.26
4	HBUT	YOLOv4 (ours)	98	98.84	99.08	91

the defects detection of insulators and provides a competitive solution for the intelligent inspection of the power grid.

The detailed cross-sectional comparisons between our model and other previous works are shown in Table 2. In this table, three insulator defect detection models with the best results are selected and name as No.1–No.3, respectively. Specifically, No.1 refers to the defect detection algorithm based on YOLOv5 from Hefei Institutes of Physical Science

[15]. According to the paper, its mAP can reach 99.05% and its accuracy can reach 86.8%. No.2 is the algorithm from Zhengzhou University [21], which incorporates Faster R-CNN network and Resnet-101 network, with an accuracy of 96.83%. No.3 represents the algorithm from Wuhan University, which has a relatively high mAP and accuracy of 93.81% and 97.26%. No.4 is the insulator defect detection algorithm based on YOLOv4 proposed in this paper.

To further verify the generality of the proposed method, experiments are performed with the VisDrone 2019 dataset [32][33]. Recognizing objects in UAV images has been a hot topic in the field of object detection. The VisDrone 2019 dataset was collected by the AISKYEYE team at the machine learning and data mining lab of Tianjin University and used as the competition dataset at the international conference on computer vision (ICCV) 2019 workshop. It contains 8599 images, including 6471 images as training set, 548 images as validation set, and 1610 images as test set. There are a total of 10 categories in this dataset, including van, bus, person, truck, car, awning tricycle, bicycles, pedestrians, cars, and tricycles.

For a more comprehensive comparison, a series of work have been done and presented in Table 3. The detection results of our improved method are compared with that of YOLOv3 and YOLOv4-tiny. Specifically, the first column of Table 3 represents the network name, the next four columns enumerate the AP values of typical categories in the dataset, including car, bus, van and truck, the last column is the mAP value. The best detection mAP of 10.68 is reported by our proposed method, while the AP values in the three items of bus, van and truck are also higher than YOLOv3 and YOLOv4-tiny. In conclusion, the proposed algorithm has better detection performance on the VisDrone dataset.

3.2 Effect of Improved Methods

A series of comparison experiments have been conducted to evaluate the effectiveness of the proposed method. The detection results of yolov4 algorithm trained by different

Table 3 YOLOv4 versus YOLOv3 performance for VisDrone

Model	AP				mAP (%)
	Car (%)	Bus (%)	Van (%)	Truck (%)	
YOLOv3	43	25	7.4	7.7	9.38
YOLOv4-tiny	11	12	2	4	3.04
YOLOv4(ours)	40	30	13	10	10.68

Bold values are the best results of the evaluation indicators

Table 4 The effect of data augmentation and anchor box redesign

Model	Training set	Testing set	Anchor box size	Precision (%)
YOLOv4	763	85	(12,16), (19,36), (40,28), (36,75), (76,55), (72,146), (142,110), (192,243), (459,401)	62.10
YOLOv4 after data augmentation	2544	848	(12,16), (19,36), (40,28), (36,75), (76,55), (72,146), (142,110), (192,243), (459,401)	75.15
YOLOv4 after anchor redesign	763	85	(23,22), (86,21), (112,35), (186,46), (251,180), (279,87), (291,47), (293,65), (294,121)	52.19
YOLOv4 with anchor redesign and data augmentation	2544	848	(23,22), (86,21), (112,35), (186,46), (251,180), (279,87), (291,47), (293,65), (294,121)	99.20

improved methods are shown in Table 4. The first line represents training with the original method, the second line represents training after data augmentation, the third line represents training after redesigning the size of the anchor box and the last line represents training after data expansion and anchor box redesign at the same time.

It can be found that the precision decreases slightly when training with new anchor boxes, which may be due to insufficient samples. The precision value is significantly improved after data enhancement. Using data enhancement and K-means method, the precision is greatly improved. It is proved that both methods can improve the performance of the experiment.

4 Discussion

4.1 Influence of Features on Training

In this section, the effects of different features on training, including cosine annealing scheduler, label smoothing, and mosaic data augmentation, are introduced and analyzed.

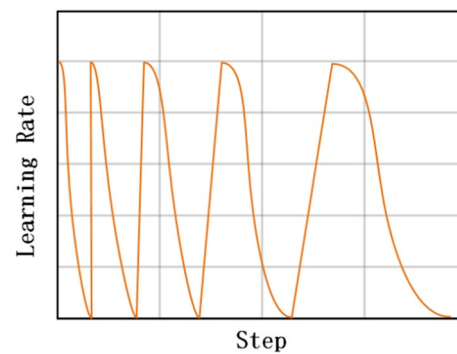


Fig. 9 Schematic diagram of updating learning rate based on cosine annealing scheduler

Table 5 Training network under different feature combinations

Experiment number	Label smoothing	Cosine annealing	Data enhancement method
Strategy 1			
Strategy 2			✓
Strategy 3		✓	✓
Strategy 4	✓	✓	✓

Table 6 Experimental results of different strategies

Condition	mAP (%)			
	Strategy1	Strategy2	Strategy3	Strategy4
IoU @0.25	74.56	99.59	72.98	70.77
IoU @0.50	73.44	99.08	71.16	69.49
IoU @0.75	41.75	58.32	58.32	58.32
IoU @0.90	5.13	6.53	6.53	0.03

4.1.1 Altering Learning Rate with the Cosine Annealing Scheduler

The learning rate is updated by the cosine annealing scheduler to simulate the learning restart process. It starts with a large learning rate, drops to the minimum value relatively quickly, and then increases rapidly to avoid the weight falling into local minimum value during the gradient descent process. A schematic diagram of updating the learning rate

based on a cosine annealing scheduler is shown in Fig. 9. In the figure, the vertical axis represents the learning rate and the horizontal axis corresponds to the learning process, which is a covariate directly related to time.

4.1.2 Label Smoothing

In addition, to prevent the network from overfitting and improve its generalization ability during training, label smoothing will be used to convert hard labels to soft labels. It is a regularization method, that adds penalty factors to refine the output labels. The implementation process of the method is shown as:

$$\hat{y}_i = y_{hot}(1 - \alpha) + \alpha/num \tag{7}$$

where the \hat{y}_i represents the redesigned label, y_{hot} represents the original label, α represents the value of the smooth label, num represents the number of categories. In this paper, the value of label smoothing is set to 0.1.

The network is trained with different strategies, of which each strategy is tested with different IoU values. The details of the training strategy are shown in Table 5, the mAP values of the detection result are shown in Table 6. The IoU@0.25, IoU@0.50, IoU@0.75, and IoU@0.90 represent the map values of the network prediction results when IoU is set to 0.25, 0.5, 0.75, and 0.9, respectively. It can be seen from Table 5 that the accuracy of the algorithm decreases as the IoU threshold increases. For these four strategies, when the IoU value is set to 0.9, the value of mAP goes down to the

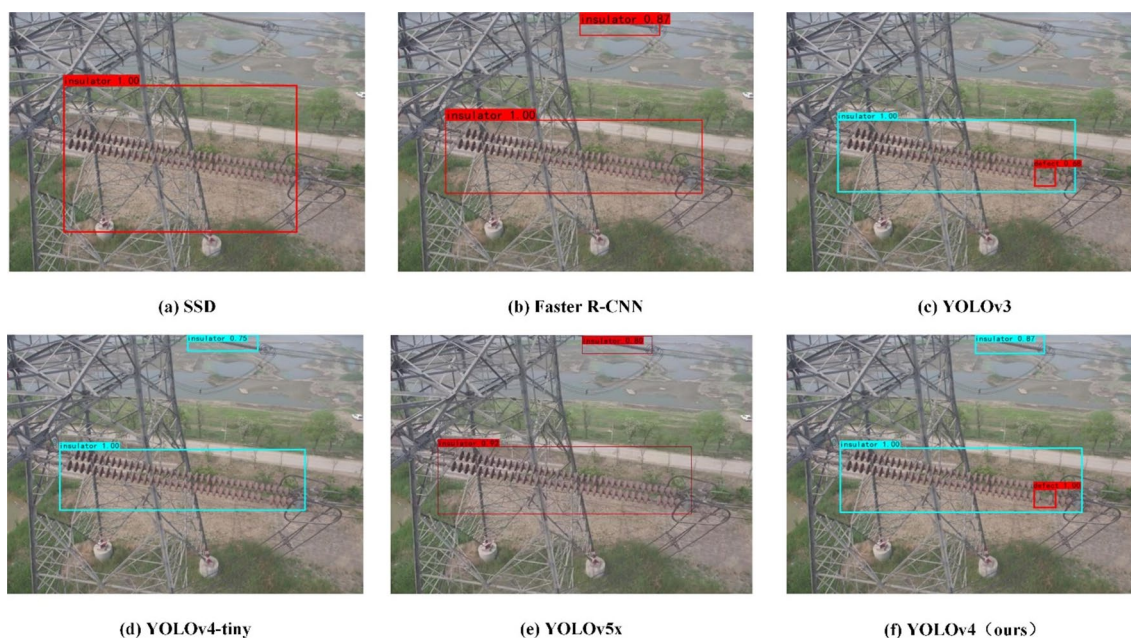


Fig. 10 The detection result under good light conditions

minimum. The reason for this phenomenon is if the IoU value is set too large, some prediction objects with low confidence will be filtered out, thus missing the network prediction. When the value of IoU is set to 0.5 or 0.75, the mAP value of strategy 2 is significantly higher than that of other strategies. Based on the above experiments results, for the purpose of improving the recall rate of prediction results, the value of IoU is set to a relatively lower value of 0.5 in this paper.

5 Robustness Test

Some typical images with different light intensities and different background complexity are detected, and they are compared with other models to further endorse the detection effectiveness of the proposed method.

The results are shown in Figs. 10, 11, and 12, respectively.

Figure 10 shows a comparison of the detection results under good light conditions. It can be seen that all objects in the image are recognized by YOLOv4 in Fig. 10f, and the predicted box matches the actual box best. In Fig. 10a, the inspection of the defects and the distant insulator are missed. In Fig. 10c the defect is detected by YOLOv3 but the insulator in the distance is also missed. The Faster R-CNN, YOLOv4-tiny and YOLOv5x accurately located the insulator, but all miss the defects in Fig. 10b, e and d.

Figure 11 shows the detection result of an image with poor light conditions. As shown in Fig. 11f, the proposed

network still accurately located all the objects. In Fig. 11c, defects are detected by YOLOv4-tiny, but the location is not accurate enough, and the bounding box of defect is larger than the actual box. However, the other algorithms as shown in Fig. 11a, b, c and e only locate the insulators and miss the defects. Both of these two algorithms are difficult to withstand the influence of changes in the outdoor detection environment.

Figure 12 shows the detection comparison of different algorithms under complex backgrounds. Because the shooting angle cannot be precisely controlled, part of the insulators in the images are blocked. As shown in Fig. 12a, b, c and d, the insulators blocked by the transmission line pylon are missed. In Fig. 12f, all objects are detected using the YOLOv4(ours). It has high robustness and high detection accuracy, thus it is more suitable for insulator defect detection.

6 Conclusions

This paper discusses the insulator defects detection of the power grid aerial images. Firstly, the YOLOv4 network is used as the basic model to analyze the defect feature and optimize the size of anchor boxes. Secondly, in response to the problem of fewer data sets, the dataset is expanded by the proposed augmentation method. Finally, under the same experimental conditions, the proposed method is compared

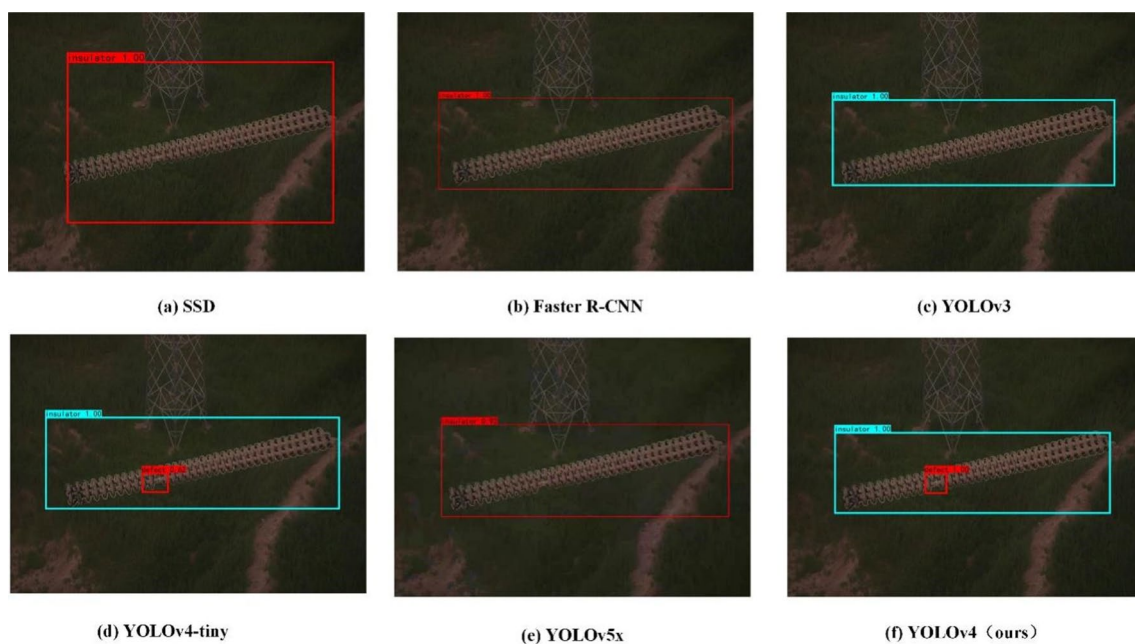


Fig. 11 The detection result under the poor light conditions

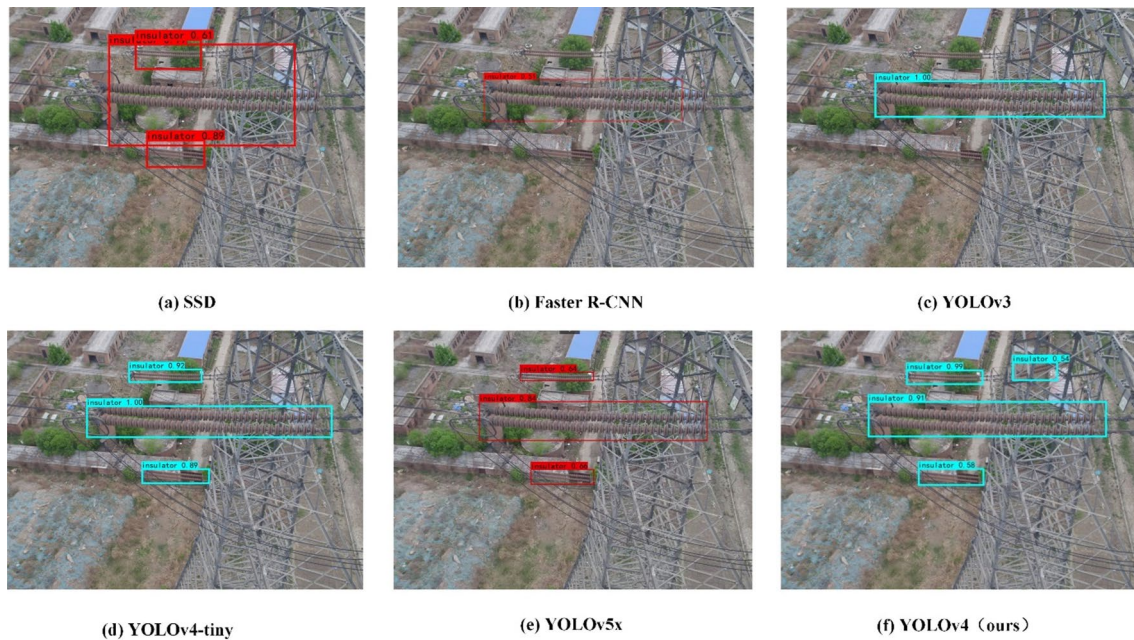


Fig. 12 The detection result under complex background.

with other advanced algorithms. This paper draws the following conclusions from the experiments:

- (1) By implementing data enhancement and anchor boxes redesign, our proposed method is superior to other comparison algorithms, including SSD, Faster-RCNN and released version of YOLO. Compared to the streamlined version of the YOLOv4, our algorithm improves the detection precision by 37.2%. Moreover, it also outperforms YOLOv5x across the board in several evaluation metrics such as Recall, AP value, mAP and F1 of insulator defect detection results.
- (2) We also conduct a cross-sectional comparison of our proposed method with other previous works and further validate its generality on the VisDrone 2019 dataset. The simulation results all show the effectiveness and superiority of the proposed method in this paper.
- (3) The robustness test results demonstrate that our proposed method performs well under different light intensities and complex environmental backgrounds, and can accurately detect all targets, which is significantly better than other comparative algorithms.

However, the inspection task of the power transmission systems is not only the defect detection of the insulators but also the common components that may cause the power grid failure including poles, ground wires, fittings, etc. In the future, multi-fault detection and classification will be carried out at the same time, and visual intelligent detection

software will be constructed to realize intelligent detection of power transmission system.

7 Supporting information

Images and data from this study are available on Figshare at: https://figshare.com/articles/dataset/Power_Line_Insulators_Dataset/11826483<https://doi.org/10.5281/zenodo.3656611>.

Acknowledgments The author would like to thank Raimundo, António, Sebastião, et al. for making the CPLID publicly available.

Funding This research is supported by Open Foundation of Hubei Engineering Research Center for Safety Monitoring of New Energy and Power Grid Equipment (HBSKF202127), the Key Project of Science and Technology Research Program of Hubei Educational Committee (D20211402) and Open Foundation of Hubei Key Laboratory for High-efficiency Utilization of Solar Energy and Operation Control of Energy Storage System (HBSEES202106).

Declarations

Conflict of interest The authors declare that there is no conflict of interest.

References

1. Zhao W, Xu M, Cheng X et al (2021) An insulator in transmission lines recognition and fault detection model based on improved Faster RCNN. *IEEE Trans Instrum Meas* 70:1–8. <https://doi.org/10.1109/TIM.2021.3112227>

2. Yang Y, Wang L, Wang Y et al (2019) Insulator self-shattering detection: a deep convolutional neural network approach. *Multimed Tools Appl* 78(8):10097–10112. <https://doi.org/10.1007/s11042-018-6610-4>
3. Yang L, Fan J, Liu Y et al (2020) A review on state-of-the-art power line inspection techniques. *IEEE Trans Instrum Meas* 69(12):9350–9365. <https://doi.org/10.1109/TIM.2020.3031194>
4. Boufares F, Doudech I, Bahrami MR (2022), May electrical transmission lines robot inspector: design challenges. In: 2022 International conference on industrial engineering, applications and manufacturing (ICIEAM), 746–754. <https://doi.org/10.1109/ICIEAM54945.2022.9787220>
5. Wang C, Pei H, Tang G et al (2022) Pointer meter recognition in UAV inspection of overhead transmission lines. *Energy Rep* 8:243–250. <https://doi.org/10.1016/j.egy.2022.02.108>
6. Liu Z, Wu G, He W et al (2022) Key target and defect detection of high-voltage power transmission lines with deep learning. *Int J Electr Power Energy Syst* 142:108277. <https://doi.org/10.1016/j.ijepes.2022.108277>
7. Ren Y, Tsai MY, Chen L et al (2020) Measurement and control of an incubator temperature by using conventional methods and fiber Bragg grating (FBG) based temperature sensors. *J Med Syst* 44(10):1–13
8. Burunkaya M, Yucel M (2020) Measurement and control of an incubator temperature by using conventional methods and fiber Bragg grating (FBG) based temperature sensors. *J Med Syst* 44(10):1–13. <https://doi.org/10.1007/s10916-020-01650-2>
9. Samonte MJC, Garcia AC, Gorre JEE et al (2022) CrowdSurge: a crowd density monitoring solution using smart video surveillance with security vulnerability assessment. *Archit (CUDA)* 7:9
10. Tian Y, Yang G, Wang Z et al (2020) Instance segmentation of apple flowers using the improved mask R-CNN model. *Biosyst Eng* 193:264–278. <https://doi.org/10.1016/j.biosystemseng.2020.03.008>
11. Zhai Y, Chen R, Yang Q et al (2018) Insulator fault detection based on spatial morphological features of aerial images. *IEEE Access* 6:35316–35326. <https://doi.org/10.1109/ACCESS.2018.2846293>
12. Xu J, Cao J, Yang K (2016) Design of insulator image segmentation algorithm based on local average. *Comput Program* 42(9):262–267. <https://doi.org/10.3969/j.issn.1000-3428.2016.09.046>
13. Tan P, Li X, Xu J et al (2020) Catenary insulator defect detection based on contour features and gray similarity matching. *J Zhejiang Univ-Sci A* 21(1):64–73. <https://doi.org/10.1631/jzus.A1900341>
14. Wei Z (2022) Composite insulator defect identification and quantitative method based on random hough transform ellipse detection. *J Phys Conf Ser* 2170(1):012021
15. Jiang C, Wang Z, Liang H et al (2022) A fast and high-performance object proposal method for vision sensors: application to object detection. *IEEE Sens J* 22(10):9543–9557. <https://doi.org/10.1109/JSEN.2022.3155232>
16. Zhao Z, Liu Y, Sun X et al (2021) Composited FishNet: fish detection and species recognition from low-quality underwater videos. *IEEE Trans Image Process* 30:4719–4734. <https://doi.org/10.1109/TIP.2021.3074738>
17. Ren S, He K, Girshick R et al (2015) Faster R-CNN: towards real-time object detection with region proposal networks. *Adv Neural Inf Process Syst*. <https://doi.org/10.1109/TPAMI.2016.2577031>
18. Liu W, Anguelov D, Erhan D et al (2016) SSD: single shot Multi-Box detector. Springer International Publishing, New York
19. Redmon J, Divvala S, Girshick R et al (2016) You only look once: Unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779–788. <https://doi.org/10.1109/CVPR.2016.91>
20. Dian S, Zhong X, Zhong Y (2022) Faster R-Transformer: an efficient method for insulator detection in complex aerial environments. *Measurement*. <https://doi.org/10.1016/j.measurement.2022.111238>
21. Gao J, Yihang L (2019) Research on recognition and segmentation of insulator strings in aerial images. *J Zhengzhou Univ* 51(4):16–22. http://en.cnki.com.cn/article_en/cjfdtotal-zzdz201904003.htm
22. Qiu Z, Zhu X, Liao C et al (2022) Detection of transmission line insulator defects based on an improved lightweight YOLOv4 model. *Appl Sci* 12(3):1207. <https://doi.org/10.3390/app12031207>
23. Bochkovskiy A, Wang CY, Liao HYM (2020) Yolov4: optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934. <https://doi.org/10.48550/arXiv.2004.10934>
24. Tao X, Zhang D, Wang Z et al (2018) Detection of power line insulator defects using aerial images analyzed with convolutional neural networks. *IEEE Trans Syst Man Cybern Syst* 50(4):1486–1498. <https://doi.org/10.1109/TSMC.2018.2871750>
25. Redmon J, Farhadi A (2018) Yolov3: An incremental improvement. arXiv preprint arXiv. <https://doi.org/10.48550/arXiv.1804.02767>
26. Wang CY, Liao HYM, Wu YH et al (2020) CSPNet: A new backbone that can enhance learning capability of CNN. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp 390–391. <https://doi.org/10.1109/CVPRW50498.2020.00203>
27. He K, Zhang X, Ren S et al (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans Pattern Anal Mach Intell* 37(9):1904–1916. <https://doi.org/10.1109/TPAMI.2015.2389824>
28. Lin TY, Dollár P, Girshick R et al (2017) Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp 2117–2125. <https://doi.org/10.1109/CVPR.2017.106>
29. Hanchao L, Pengfei X, Jie A, Lingxue W et al (2018) Pyramid attention network for semantic segmentation. arXiv preprint arXiv: <https://doi.org/10.48550/arXiv.1805.10180>
30. Joseph R, Ali F (2018) Yolov3: an incremental improvement. arXiv preprint arXiv:180402767. <https://doi.org/10.48550/arXiv.1804.02767>
31. Zhan W, Sun C, Wang M et al (2022) An improved Yolov5 real-time detection method for small objects captured by UAV. *Soft Comput* 26:361–373. <https://doi.org/10.1007/s00500-021-06407-8>
32. Pengfei Z, Longyin W, Xiao B et al (2018) Vision meets drones: a challenge. arXiv preprint arXiv: 1804.07437. <https://doi.org/10.48550/arXiv.1804.07437>
33. Pengfei Z, Longyin W, Dawei D et al (2020) Vision meets drones: past, present and future. arXiv preprint arXiv: 2001.06303. <https://doi.org/10.48550/arXiv.2001.06303>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Shanshan Wang an Associate Professor and Master Tutor of School of Electrical and Electronic Engineering, Hubei University of Technology. She received the B.S. degree from the Huazhong University of Science and Technology in 2003, the M.S. degree from the Huazhong University of Science and Technology in 2006, and the Ph.D. degree from the Huazhong University of Science and Technology in 2018.

From 2006 to 2007, she has been a senior engineer with ZTE Communications Co., Ltd. From 2008 to 2014, she has been an associate professor with the Department of Information Engineering, Wuchang Institute of Technology. Since 2018, she has been an associate professor with Hubei University of Technology. Her research interests include machine learning and deep learning, medical image analysis, ultrasound imaging, image processing, etc.

Xinyi Zou a graduate student of School of Electrical and Electronic Engineering, Hubei University of Technology. She received the B.S. degree from the Wuchang Institute of Technology, Wuhan, China, in 2019, the M.S. degree from the Hubei University of Technology, Wuhan, China, in 2022. His research interest covers machine learning and deep learning, image processing, object detection, etc.

Wei Zhu is a graduate student of the School of Electrical and Electronic Engineering, Hubei University of Technology. He received his bachelor's degree in engineering from Hubei University of Technology in 2020 and has studied at Hubei University of Technology since

2020. His research interests include deep learning and medical image classification.

Liang Zeng a Professor and Master Tutor of School of Electrical and Electronic Engineering, Hubei University of Technology. He also is a deputy secretary-general and executive director of the Hubei Society of Artificial Intelligence. He graduated from Wuhan University of Technology with a bachelor's degree of industrial automation in 2001 and a master's degree in control theory and control engineering in 2004, respectively. In 2008, he graduated from Huazhong University of Science and Technology with a PhD in Systems Engineering. From June 2008 to March 2017, he worked as a senior researcher at the Technology Research Institute of WISDRI Engineering & Research Incorporation Limited. His research interests include artificial intelligence, evolutionary algorithms and their applications, modeling and optimization of complex systems, etc. Up till now, he has published more than 20 academic papers.