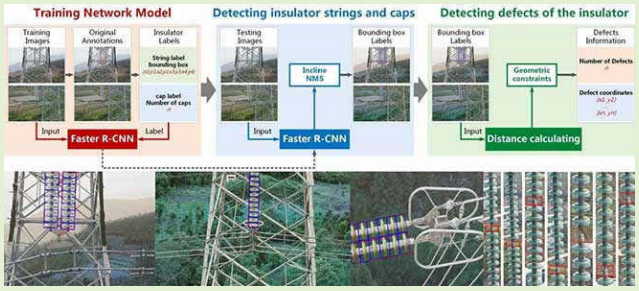
Cap-Count Guided Weakly Supervised Insulator Cap Missing Detection in Aerial Images

Chunxue Shi and Yaping Huang

***Abstract*—The vision-based detection of cap missing on insulator in aerial images is an extremely challenging task. Previous approaches have achieved promising performances. However, most of them include two stages of insulator localization and cap detection, which require large-scale insulator and cap annotations. In this article, we propose a cap-count guided weakly supervised approach, which only needs insulator bounding box annotations. The proposed method first generates cap annotations automatically by giving the insulator bounding boxes and the number of caps. Then, we use a deep learning model to detect insulator strings and caps simultaneously. Finally, we design an efficient cap missing detection strategy by using caps geometric constraints to measure the position of caps. We evaluate the proposed approach on our collected large-scale aerial image dataset. The experimental results show that the proposed method can accurately detect the insulator strings and caps with any angle from an aerial image and further recognize the missing cap on insulator. The precision and recall reach 92.86% and 86.67% respectively, which achieves competitive detection results.**



***Index Terms*— Aerial image, insulator, cap, deep learning, cap missing detection.**

I. Introduction

INSULATORS are important components on high-voltage power lines for increasing shooting distance and securing wires. In harsh natural environments, insulators on overhead transmission lines will experience aging, corrosion, damage and other failures. The failure of the insulator is a serious threat to the normal operation of the power system. The defects of the insulator is mainly caused by the self-detonation of the cap, which resulting in the missing of the insulator caps. In recent years, unmanned aerial vehicle (UAV) detection becomes a common way of detecting transmission lines. However, it is a challenging task to detect the missing of cap on insulator due to small defect regions and complex backgrounds.

A. Related Works and Challenges

A number of insulator detection approaches have been proposed to deal with the detection of the missing insulator cap. Object detection-based methods and multi-stage methods are the two mainstream directions of insulator detection tasks.

Manuscript received January 16, 2020; revised June 1, 2020; accepted July 20, 2020. Date of publication July 29, 2020; date of current version December 4, 2020. This work was supported by the Fundamental Research Funds for the Central Universities under Grant 2019JBZ104. The associate editor coordinating the review of this article and approving it for publication was Dr. Thilo Sauter. *(Corresponding author: Yaping Huang.)*

The authors are with the Beijing Key Laboratory of Traffic Data Analysis and Mining, Beijing Jiaotong University, Beijing 100044, China (e-mail: yphuang@bjtu.edu.cn; 17120408@bjtu.edu.cn).

Digital Object Identifier 10.1109/JSEN.2020.3012780

1) Object Detection-Based Methods: Benefiting from the progress of deep learning technologies, object detection-based methods can directly localize the defect regions in aerial images. Liu *et al.* [1] proposed to detect defects in insulator based on Faster R-CNN, which regards the cap missing region as a target region. Tao *et al.* [2] proposed a novel deep convolutional neural network (CNN) cascading architecture for performing localization and detecting defects in insulators. Liao *et al.* [3] proposed an effective insulator defect detection model using improved Faster Region-based Convolutional neural network (Faster R-CNN) based on the original aerial insulator image. The soft non-maximum suppression (Soft-NMS) is utilized to improve the performance of detection overlap insulator and ResNet 101 model is adopted to effectively extract features from insulator images.

However, there are few insulator defect samples in real data, so above methods suffer from lack of training samples and the interference of complex backgrounds.

2) Multi-Stage Methods: Multi-stage methods are widely used in insulator detection. Most of them include insulator localization, cap detection and defect localization. For insulator string detection, Wang *et al.* [4] proposed to combine the shape, color and texture feature of insulator to detect insulators. Wang *et al.* [5] put forward to detect the self-detonation defects of insulators via target detection, segmentation and classification. Similarly, Gao *et al.* [6] utilized a convolutional neural network model to achieve detection of insulators. Zhao *et al.* [7] proposed an insulator localization method

based on orientation angle detection and binary shape prior knowledge (OAD-BSPK). Literature [8], [9] proposed to learn the characteristics of insulators in complex aerial images through convolutional neural networks, and then identify various insulators. Tiantian *et al.* [10] used fused feature to locate multiple insulators with different angles based on threshold segmentation. Zuo *et al.* [11] proposed a two-stage method. It first obtains a classifier identifying and locating the insulator by feature extraction and training, and then segments the insulator region through a series of digital image processing methods. For insulator defects detection, Miao *et al.* [12] proposed an effective and reliable insulator detection method based on deep learning techniques. Zhai *et al.* [13] designed a bunch-drop fault detection method for both glass and ceramic insulators based on spatial morphological features. It is able to locate the ceramic insulators and glass insulators, but can not automatically distinguish the type of the insulator.

Recently, some new ideas are also proposed for detecting insulators. Sun *et al.* [14] proposed an insulator identification and localization method based on three-dimensional point cloud modeling. It extracts the depth information from three-dimensional point cloud data, and identifies the insulator based on principal component analysis to obtain the position of the insulator in three-dimensional space.

Although multi-stage methods achieve promising results, however, those methods can not tightly locate insulators with different directions. Moreover, they require expensive cost to obtain the pixel-level annotations of the insulators and caps.

In summary, the main challenge of the missing insulator caps detection lies in three points.

1) Insulator aerial images are usually captured under the complex background. However, most existing approaches use axis-aligned bounding boxes to locate the insulator strings, which may result in the serious background noise. Therefore, how to effectively eliminate the interference of the complex background is a difficult problem in the insulator detection task.
2) Due to the shooting angle, the tilt angle of the insulator string in the aerial image is different. How to use an effective method to accurately locate the insulator region, and further facilitate the location of the missing cap is the key issue of the missing cap detection.
3) Most existing approaches require laborious works to annotate the insulator strings and caps. Therefore, how to reduce the cost of annotations is worthy to be considered carefully, especially in practical scenario.

B. Our Contributions

In order to solve the problems, in this article, we propose a novel cap-count guided weakly supervised insulator detection approach to detect the missing of caps. Our proposed method not only can eliminate the interference from complex backgrounds, but also only requires insulator bounding box annotation and the corresponding cap counts (as shown in Fig. 1(b)).

The overall framework is shown in Fig. 2. Specifically, we propose to generate cap annotations automatically by employing the insulator bounding boxes and the corresponding cap counts. Our proposed approach can be regarded as a weakly-supervised method due to the fact that we only need the number of caps instead of the coordinates of each cap. It greatly reduces the workload of annotation and saves time. Then, we use (x_1, y_1, x_2, y_2, h) to define the bounding box named *HBOX*, representing the target region. The generated detection result is an inclined bounding box, so that the insulator strings in different directions can be tightly located. More importantly, we use a deep learning model to detect insulator strings and caps simultaneously. Finally, we use caps geometric constraints to measure the position of caps and locate the missing caps. The detection pipeline for insulator strings and caps is a single network, and it is possible to directly optimize the performance of the detection.

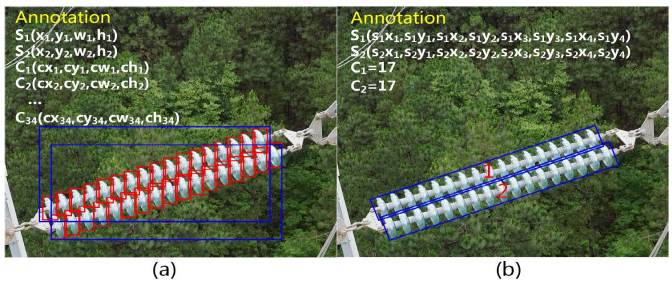


Fig. 1. Our proposed approach does not require laborious insulator and cap annotations (a). Instead, we only need to define the inclined insulator boxes and the number of caps in them (b).

We conduct extensive experiments on our collected large-scale aerial image dataset. The experimental results show that our method can accurately detect the insulator strings and caps with any angle in the aerial images and further recognize the missing cap on insulator strings.

The contribution of this article are as follows:

- We propose to generate cap annotations automatically by employing the insulator bounding boxes and the corresponding cap counts, which greatly reduces the workload of annotation.
- We propose a strategy for detecting insulator strings and caps by inclined bounding boxes simultaneously, which can detect the caps directly with the constraints of insulator regions.
- We collect a large-scale insulator dataset including 3700 aerial images and four types of insulators. Our proposed method achieves competitive results on it compared with other methods.

II. The Proposed Approach

Our proposed approach can be regarded as a weakly-supervised method due to the fact that we only need the number of caps instead of the coordinates of each cap. It consists of three parts as shown in Fig. 2: (1) we train the network with inclined bounding boxes and the number of caps to detect the insulator strings and caps, (2) we detect the inclined insulator strings and caps simultaneously, and (3) we localize the defects of the insulator according to the coordinates of the insulator string and caps. The remainder of this section describes our proposed data annotation strategy, the definition of inclined bounding box, network architecture

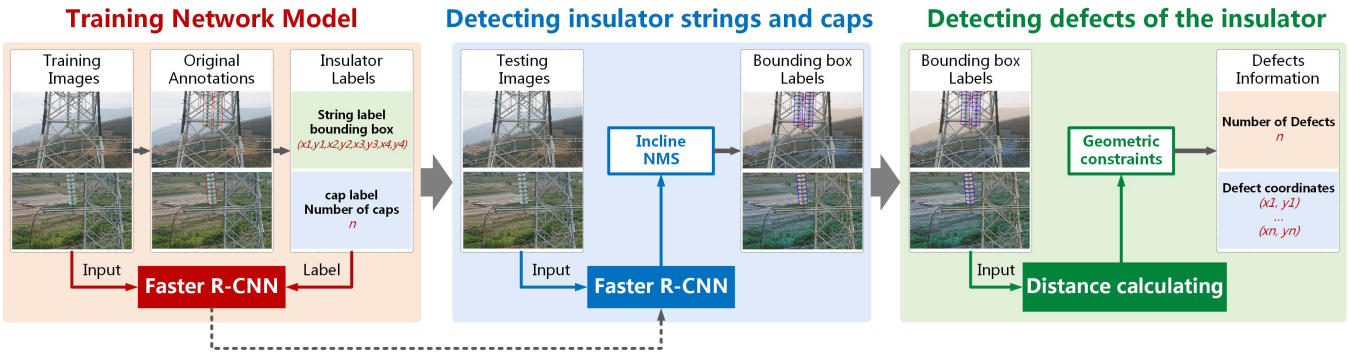


Fig. 2. The pipeline of our approach. We train the Faster R-CNN with the coordinates of the insulator strings and the number of caps to get the network model. The trained Faster R-CNN predicts inclined bounding boxes, which represent the region of insulator strings and caps. Then, we perform an inclined NMS (non-maximum suppression) on these bounding boxes to get the coordinates of insulator strings and caps. Finally, we geometrically constrain the coordinates of caps by the location of insulator string, and calculate the number and position of the missing caps on insulator.

and loss function, and the localization of insulator defects in details.

A. Data Annotation

Most of detection methods use horizontal or vertical rectangles to mark the insulator region. Fig. 3(a) shows an example image where the insulators are denoted by green boxes. However, those bounding boxes may contain other insulators and complex background, which will dramatically weaken the performance of detecting defects.

In order to tackle the problems, we propose a novel way to annotate the insulator and cap regions. We adopt the inclined bounding boxes so as to localize the insulators tightly, and we further give the number of caps within the regions to indicate the caps. Comparing with traditional methods, our proposed annotations can provide location information with an easier way, which not only alleviate expensive and laborious labelling works, but also avoid the interference of the background by introducing inclined bounding boxes.

Specifically, for one insulator string in an image, the annotation is given by the coordinates of the four vertices of the insulator string $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$ and the number of the caps n within the insulator region. We specify that the first point (x_1, y_1) always means the left-top corner of the insulator string, and the remaining coordinates are arranged in a clockwise direction, as shown in Fig. 3(b). Obviously, it is much easier for annotators to label the insulator and cap regions.

Furthermore, for training the models, we need to simply transform the original annotations to the form of $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$ for each insulator and cap. We develop a labeling tool to perform the transformation task. The algorithm used in the labeling tool is shown in Algorithm 1, where the input of the algorithm is the coordinates of the four vertices of the insulator string $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$ and the number of the caps n within the insulator region, and the output includes the coordinates of the insulator string and each cap.

B. Inclined Bounding Box Definition

In fact, the inclined bounding box can cover most of the insulator regions and eliminate the influence of the background region on the detection results. Therefore, we consider the insulator detection task with any direction by solving the problem of detecting the smallest oblique rectangular region.

Algorithm 1 Labeling Caps

Input: The coordinates of insulator string, x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4; The number of insulator caps, n;
Output: The coordinates of insulator caps in set S;

function GENERATE CAP LABELS(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, n)
  $S \leftarrow \emptyset$, Cx_1, Cy_1, Cx_2, Cy_2, Cx_3, Cy_3, Cx_4, Cy_4
  $i \leftarrow 0$
  while $i < n$ **do**
    $Cx_1 \leftarrow x_4 + (x_3 - x_4)/n * i$
    $Cy_1 \leftarrow y_4 + (y_3 - y_4)/n * i$
    $Cx_2 \leftarrow x_1 + (x_2 - x_1)/n * i$
    $Cy_2 \leftarrow y_1 + (y_2 - y_1)/n * i$
    $Cx_3 \leftarrow x_1 + (x_2 - x_1)/n * (i + 1)$
    $Cy_3 \leftarrow y_1 + (y_2 - y_1)/n * (i + 1)$
    $Cx_4 \leftarrow x_4 + (x_3 - x_4)/n * (i + 1)$
    $Cy_4 \leftarrow y_4 + (y_3 - y_4)/n * (i + 1)$
    $C \leftarrow \{Cx_1, Cy_1, Cx_2, Cy_2, Cx_3, Cy_3, Cx_4, Cy_4\}$
    $S \leftarrow S \cup C$
    $i \leftarrow i + 1$
  end while
  return S
end function

In order to reduce the number of parameters in the model training, we convert the representations of $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$ into (x_1, y_1, x_2, y_2, h) for inclined bounding box regression, as shown in Fig. 3(c).

It is very important to evaluate the distance between two boxes in object detection task. The common method for calculating the distance between two standard rectangular boxes is to calculate the IoU. In our scenario, we also choose IoU to measure the distance between two inclined boxes, A and B, as follows:

$$IoU(A, B) = \frac{area(A \bigcup B)}{area(A \bigcup B)}, \tag{1}$$

where the intersection and the union are calculated using the shared monte-Carlo method [15].

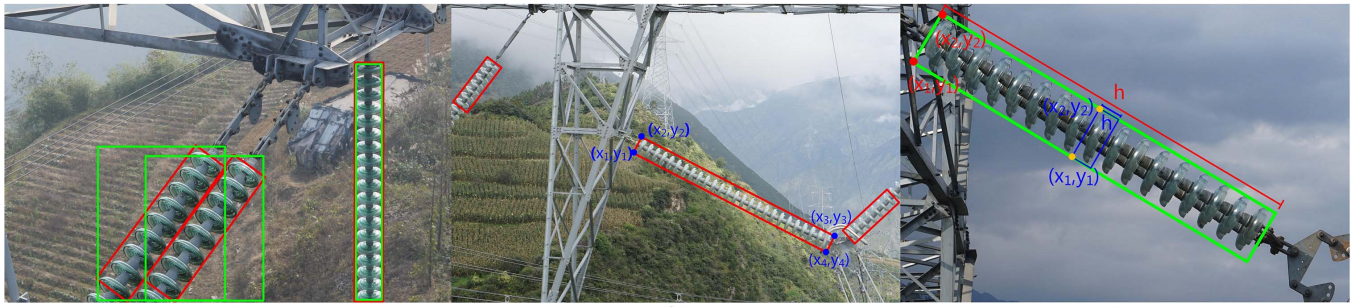


Fig. 3. (a) The annotation of insulator strings in aerial image. (b) Our proposed annotation method for insulator strings. (c) The annotation of insulator string and caps in training network model.

C. Network Architecture

We use a popular two-stage object detection model Faster R-CNN [16], including regional proposal network (RPN) and regional classification, as the backbone model. The RPN is first used to generate the insulator region proposals, which are axis-aligned bounding boxes that enclose the arbitrary-oriented insulator strings and caps. Then for each proposal, several ROI Poolings with different pool sizes (7×7, 11×3, 3×11) are performed on the convolutional feature maps and the pooled features are concatenated for further classification and regression. Subsequently, we predict the score of the target area/non-target area, axis-aligned boxes and inclined minimum area boxes. Finally, the inclined boxes are post-processed by inclined non-maximum suppression to obtain the detection results.

1) Inclined Bounding Boxes Generation: According to the geometric characteristics of the insulator string, we use RPN to generate an axis-aligned bounding box that surrounds the insulators at any angle. Specifically, we use a smaller anchor ratio in the RPN to detect smaller insulator caps. The ratio of the original anchor in Faster R-CNN [16] is (8,16,32), But we change the anchor scale to (4,8,16) so as to benefit for detecting small insulator caps. After RPN, we classify the proposals as target or non-target, then refine the axis-aligned bounding boxes that contain the arbitrary-oriented insulators and predict inclined bounding boxes. It is noted that each inclined box is associated with an axis-aligned box.

2) ROI Pooling of Different Pooled Sizes: The Faster R-CNN framework performs ROI Pooling on the feature map. The proposed pooling size for each RPN proposal is 7×7. In general, the length of the insulator strings is much larger than their width. Therefore, we propose a multi-scale pooling strategy to capture more distinctive features for insulators and their caps. Specifically, we add two pool sizes: 11×3 and 3×11. Acturally, we observe that 3×11 can capture horizontal features to detect horizontal insulators and 11×3 can capture vertical features for detecting vertical insulators. Then, we perform an inclined NMS and replace the traditional IoU calculation with the IoU between the two inclined bounding boxes.

3) Loss Function of the Network: The loss function defined for each image is the sum of the classification loss and the bounding box regression loss. The multi-task loss function is defined as follows:

$$L = L_{cls} + \lambda_1 \sum_i L_{reg\_align} + \lambda_2 \sum_i L_{reg\_incline}, \quad (2)$$

where i is the index of $i\text{-}th$ bounding box, and λ_1 and λ_2 are parameters that control the balance between three terms. The first term defines the classification loss. The second and the third term specify the regression loss, where the difference between them is that the second is the loss of the axis-aligned bounding box and the third is the loss of the minimum inclined bounding box. The L_{cls} is defined as follows:

$$L_{cls} = -\frac{1}{N} \sum_{i=0}^{N} p_i^* log p_i, \quad (3)$$

where p is the predicted probability of $i\text{-}th$ bounding box being an object, and p^* indicates the ground-truth. N is the total number of the anchor boxes. In our article, there are three class categories, representing the background, strings and caps, respectively.

For axis-aligned bounding box regression, we define the L_{reg_align} as follows:

$$L_{reg\_align} = smooth_{L1}(v_i - v_i^*), \quad (4)$$

in which

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & if\ |x| < 1 \\ |x| - 0.5 & otherwise, \end{cases} \quad (5)$$

where v_i is a vector representing the four parameterized coordinates of the predicted axis-aligned bounding box, and v_i^* is the coordinates of the smallest bounding rectangle of the inclined ground-truth box. Concretely, we adopt the parameterizations of the four coordinates as following:

$$\begin{aligned} v_x &= (x - x_a)/w_a, \quad v_y = (y - y_a)/h_a, \\ v_w &= log(w/w_a), \quad v_h = (h/h_a), \\ v_x^* &= min(x_1, x_2, x_3, x_4) + max(x_1, x_2, x_3, x_4)/2, \\ v_y^* &= min(y_1, y_2, y_3, y_4) + max(y_1, y_2, y_3, y_4)/2, \\ v_w^* &= max(x_1, x_2, x_3, x_4) - min(x_1, x_2, x_3, x_4), \\ v_h^* &= max(y_1, y_2, y_3, y_4) - min(y_1, y_2, y_3, y_4), \end{aligned} \quad (6)$$

where x, y, w, and h denote the center coordinates, width and height of the axis-aligned bounding box respectively. x and x_a represent the predicted box and the anchor box respectively

Fig. 4. The ground-truth of insulators including four categories.

(likewise for y, w, h). Note that $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$ is the coordinate of the inclined bounding box as mentioned in the section II.A.

Similarly, the $L_{reg\_incline}$ can be formulated as

$$L_{reg\_incline} = smooth_{L1}(u_i - u_i^*), \tag{7}$$

where $u_i = (u_{x_1}, u_{y_1}, u_{x_2}, u_{y_2}, u_h)$ represents the five parameterized coordinates of the predicted inclined bounding box. Specifically, (u_{x_1}, u_{y_1}) is the coordinates of the left-top point, (u_{x_2}, u_{y_2}) is the coordinates of the right-top point, and u_h is the corresponding height. $u_i^* = (u_{x_1}^*, u_{y_1}^*, u_{x_2}^*, u_{y_2}^*, u_h^*)$ indicates the ground-truth box of insulator string or cap regions.

D. Localization of Missing Caps on Insulator

After training the network, we can localize the insulator strings and caps with inclined bounding boxes, then find the cap regions with the constraint of the insulator string regions. Specifically, we only count the caps in the bounding box of the insulator string to detect the missing caps. Assuming that each insulator string contains n caps denoted as c_0, c_1, …, c_{n-1}. The distances between the center points of adjacent caps are calculated by the coordinate of each cap, which are denoted as d_0, d_1, …, d_{n-2}. We calculate the mean of the distance which is defined as

$$d_m = (d_0 + d_1 + \ldots + d_{n-2})/(n-1). \tag{8}$$

Then we determine if the cap is missing by the distances d_i $(i = 0, \ldots, n-2)$ and threshold t, which is defined as $t = \lambda \times d_m$. We define the $\lambda = 1.5$ in this article. The equation for judging whether the cap is missing can be given as follows:

$$cap\_condition(c_i) = \begin{cases} missing & if\ d_{i-1} > t \\ normal & otherwise, \end{cases} \tag{9}$$

where $i \geq 1$. Finally, we can obtain the coordinate of the missing caps.

III. Experimental Results

Our proposed approach is implemented on Tensorflow framework. We perform the experiments on a computer with Intel Xeon E5-2683 v3, 128G main memory, and a NVIDIA Tesla K40c GPU.

TABLE I
THE DETAIL OF DATASET FIN AND GCAP

Dataset	Type	Images	Insulator/Cap	Train	Val	Test
FIN1	Glass	1600	2879	960	320	320
	Ceramic	500	902	300	100	100
	Pin	800	2771	480	160	160
	Red lacquer	800	2109	480	160	160
	Sum	**3700**	**8661**	**2220**	**740**	**740**
FIN2	Glass	1600	2879	1280	160	160
	Ceramic	500	902	400	50	50
	Pin	800	2771	640	80	80
	Red lacquer	800	2109	640	80	80
	Sum	**3700**	**8661**	**2960**	**370**	**370**
GCAP1	Glass	438	6481	350	44	44
GCAP2	Glass	438	6481	394	22	22

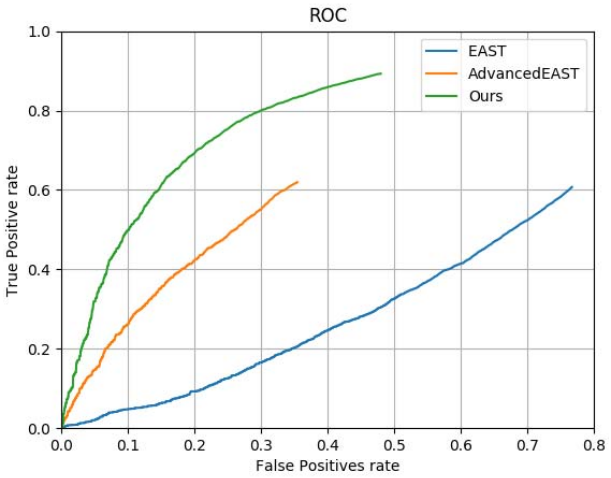


Fig. 5. The ROC curve of the insulator string detection.

A. Insulator Datasets

Our experimental data comes from the national grid, which uses drones to sample insulators on different transmission lines. We collect a large-scale dataset FIN (Four types of INsulators), which includes 3700 aerial images of UAV for insulator string detection. The aerial image size in our collected dataset is 3963 × 2624. It should be noted that our proposed method can deal with different image sizes both in training and testing stage. Specifically, we divide these samples into two datasets in two different proportions, named FIN1 and FIN2. Dataset FIN includes four types of insulators: glass, ceramic, pin and red lacquer. Moreover, we choose 438 images from FIN to construct GCAP (Glass Caps) dataset for missing caps detection. The GCAP consists of 694 insulator strings and 6481 insulator caps. Similarly, we also divided the cap samples into two datasets, named GCAP1 and GCAP2. The marked insulators are shown in Fig. 4 and the details of dataset FIN and GCAP are shown in Table I.

We use precision, recall and F1-measure to evaluate the performance of the insulator detection algorithm. The equations are as follows:

$$Precision = TP/(TP + FP), \tag{10}$$

$$Recall = TP/(TP + FN), \tag{11}$$

$$F1\text{-}measure = \frac{2 \times Precision \times Recall}{Precision + Recall}, \tag{12}$$

Fig. 6. The results of insulator detection on dataset FIN. (a) EAST (b) Advanced-EAST (c) Ours.

TABLE II
THE DETECTION RESULTS ON DATASET FIN AND GCAP

Dataset	Approaches	Precision	Recall	F1-measure
FIN1	EAST [17]	57.01%	85.34%	68.36%
	Advanced-EAST [18]	89.62%	54.91%	68.10%
	Ours	**90.05%**	**89.30%**	**89.67%**
FIN2	EAST [17]	60.53%	87.42%	71.53%
	Advanced-EAST [18]	90.41%	57.36%	70.19%
	Ours	**92.08%**	**90.45%**	**91.26%**
GCAP1	EAST [17]	90.38%	79.62%	84.66%
	Ours	**91.47%**	**93.85%**	**92.64%**
GCAP2	EAST [17]	91.03%	80.23%	85.29%
	Ours	**92.67%**	**94.04%**	**93.53%**

where TP is the number of correctly defined objects, $(TP + FP)$ is the total number of detected objects, and $(TP + FN)$ is the total number of actual objects. The precision is the ratio of the number of correctly detected targets to the total number of identified targets, and the recall is the ratio of the number of correctly detected targets to the total number of actual targets.

Besides, Receiver Operating Characteristic (ROC) curve is a commonly used evaluation metric to compare the detection performance of different methods, which illustrates the varying performance of an object detect model as its discrimination threshold is altered, so we also use ROC curve to evaluate the detection results.

B. Experimental Results and Analysis

We compare our method with EAST [17] and Advanced-EAST [18], which are two widely used methods to detect inclined objects. Meanwhile, we further evaluate the performance of cap detection with EAST [17]. The results are given in Table II. As shown in Table II, our method achieves the highest F1-measure of 91.26% on dataset FIN2, and both precision and recall are also the best results. Moreover, the ROC curve is also shown in Fig. 5. From the results, we can see that our proposed method achieves competitive results compared with other methods.

Some detection results are shown in Fig. 6. We can observe that EAST detects a long insulator string as multiple segments. The results suggest that EAST can not deal with the case of long insulator strings very well. It is because the insulator strings in aerial images have various shapes and aspect ratios, and are relatively long and narrow, which leads to the low accuracy of the EAST algorithm. Instead our method is not affected by the receptive field, and can completely detect elongated insulator strings, thereby improving accuracy than EAST by a large margin.

In addition, as shown in the fifth image of Fig. 6(b), the Advanced-EAST algorithm fails to detect two adjacent insulator strings. When the distance between multiple insulator strings is close or there is a blocking relationship, the detection result will be disordered. This is because the Advanced-EAST algorithm needs to detect the head and tail of the target when detecting the target, and then connect the head and tail according to certain rules to obtain the complete target. If the two insulator strings are close, the algorithm will connect the head of the first target with the tail of the other target. Therefore, the Advanced-EAST algorithm achieves a lower recall rate. Compared with Advanced-EAST, our method can better deal with the situation where the insulator strings are shielded from each other, so the recall rate is effectively improved.

Moreover, Advanced-EAST is suitable for text detection because there is no intersection between texts. However, in our scenario, it is difficult to handle the case where the insulator strings or caps are too close, thus we only compare our approach and EAST on cap detection. From the Table II, we can see that our approach also outperforms EAST for detecting caps on F1-measure, which suggests that our approach is more suitable in practical scenarios (some results are shown in Fig. 7(a)).

Finally, we use our method to locate the missing caps. The detection results are listed in Table III (shown in Fig. 7(b)). It can be seen that our method can accurately locate the missing caps on insulator in aerial images.

(a) (b)

Fig. 7. (a) The cap detection results on dataset GCAP. The first row is the results of the EAST and the second is the results of our method. (b) Some results of cap missing detection on dataset GCAP.

TABLE III
THE RESULTS OF CAP MISSING DETECTION ON DATASET GCAP

missing caps	Detected missing caps	Precision	Recall	F1-measure
15	13	92.86%	86.67%	90.85%

IV. CONCLUSION

In this article, we propose a cap-count guided weakly supervised approach for insulator cap missing detection. Specifically, we propose to generate cap annotations automatically. Moreover, we propose a strategy for detecting insulator strings and caps by inclined bounding boxes simultaneously. The experimental results show that our method can accurately detect the insulator strings and caps at any angle in the aerial image and further recognize the missing caps on insulator.

REFERENCES

[1] X. Liu, H. Jiang, J. Chen, J. Chen, S. Zhuang, and X. Miao, "Insulator detection in aerial images based on faster regions with convolutional neural network," in *Proc. IEEE 14th Int. Conf. Control Autom. (ICCA)*, Jun. 2018, pp. 1082–1086.

[2] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, and D. Xu, "Detection of power line insulator defects using aerial images analyzed with convolutional neural networks," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 4, pp. 1486–1498, Apr. 2020.

[3] G.-P. Liao, G.-J. Yang, W.-T. Tong, W. Gao, F.-L. Lv, and D. Gao, "Study on power line insulator defect detection via improved faster region-based convolutional neural network," in *Proc. IEEE 7th Int. Conf. Comput. Sci. Netw. Technol. (ICCSNT)*, Oct. 2019, pp. 262–266.

[4] W. Wang, Y. Wang, J. Han, and Y. Liu, "Recognition and drop-off detection of insulator based on aerial image," in *Proc. 9th Int. Symp. Comput. Intell. Design (ISCID)*, vol. 1, Dec. 2016, pp. 162–167.

[5] Y. Wang *et al.*, "Detection and recognition for fault insulator based on deep learning," in *Proc. 11th Int. Congr. Image Signal Process., Biomed. Eng. Informat. (CISP-BMEI)*, Oct. 2018, pp. 1–6.

[6] F. Gao *et al.*, "Recognition of insulator explosion based on deep learning," in *Proc. 14th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. (ICCWAMTIP)*, Dec. 2017, pp. 79–82.

[7] Z. Zhao, N. Liu, and L. Wang, "Localization of multiple insulators by orientation angle detection and binary shape prior knowledge," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 22, no. 6, pp. 3421–3428, Dec. 2015.

[8] G. Tao, C. Fengxiang, W. Wei, S. Ping, S. Lei, and C. Tianzhu, "Electric insulator detection of UAV images based on depth learning," in *Proc. 2nd Int. Conf. Power Renew. Energy (ICPRE)*, Sep. 2017, pp. 37–41.

[9] L. Ma, C. Xu, G. Zuo, B. Bo, and F. Tao, "Detection method of insulator based on faster R-CNN," in *Proc. IEEE 7th Annu. Int. Conf. CYBER Technol. Autom., Control, Intell. Syst. (CYBER)*, Jul. 2017, pp. 1410–1414.

[10] Y. Tiantian, Y. Guodong, and Y. Junzhi, "Feature fusion based insulator detection for aerial inspection," in *Proc. 36th Chin. Control Conf. (CCC)*, Jul. 2017, pp. 10972–10977.

[11] D. Zuo, H. Hu, R. Qian, and Z. Liu, "An insulator defect detection algorithm based on computer vision," in *Proc. IEEE Int. Conf. Inf. Autom. (ICIA)*, Jul. 2017, pp. 361–365.

[12] X. Miao, X. Liu, J. Chen, S. Zhuang, J. Fan, and H. Jiang, "Insulator detection in aerial images for transmission line inspection using single shot multibox detector," *IEEE Access*, vol. 7, pp. 9945–9956, 2019.

[13] Y. Zhai, R. Chen, Q. Yang, X. Li, and Z. Zhao, "Insulator fault detection based on spatial morphological features of aerial images," *IEEE Access*, vol. 6, pp. 35316–35326, 2018.

[14] Y. Sun, X. Chen, X. Jian, and Z. Xiao, "Identification and localization method of the insulator based on three-dimensional point cloud modeling," in *Proc. Chin. Control Conf. (CCC)*, Jul. 2019, pp. 7051–7056.

[15] Y. Liu and L. Jin, "Deep matching prior network: Toward tighter multi-oriented text detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3454–3461.

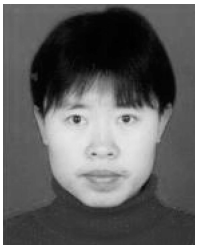
[16] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2015, pp. 91–99.

[17] X. Zhou *et al.*, "EAST: An efficient and accurate scene text detector," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2642–2651.

[18] H. Yijie. (2018). *Advanced-East*. Accessed: Sep. 4, 2018. [Online]. Available: https://github.com/huoyijie/AdvancedEAST

Chunxue Shi received the B.E. degree from the Nanjing Agricultural University of Science and Technology, Nanjing, China, in 2017. She is currently pursuing the master's degree with Beijing Jiaotong University, Beijing, China.

Her research focuses on machine learning and computer vision, especially exploring object detection in remote sensing data.



Yaping Huang received the B.E. degree from the Beijing Jiaotong University of Computer Science and Technology, Beijing, China, in 1995, and the Ph.D. degree from Beijing Jiaotong University, Beijing, in 2004.

She is currently a Professor with Beijing Jiaotong University. Her research interests include computer vision, machine learning, and pattern recognition.