

Received April 25, 2019, accepted June 9, 2019, date of publication June 14, 2019, date of current version July 2, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2923024

# Semantic Segmentation of Transmission Lines and Their Accessories Based on UAV-Taken Images

LONGJUN WANG<sup>1</sup>, ZHAOLONG CHEN<sup>1</sup>, DONG HUA<sup>1</sup>, AND ZIXUAN ZHENG<sup>2</sup>

<sup>1</sup>School of Electric Power, South China University of Technology, Guangzhou 510641, China

<sup>2</sup>International Department, Affiliated High School, South China Normal University, Guangzhou 510630, China

Corresponding author: Dong Hua (dhua@scut.edu.cn)

**ABSTRACT** While unmanned aerial vehicles (UAVs) facilitate transmission line inspection to some extent, they also generate numerous raw images rather than defect analysis results. Because analyzing the UAV-taken images by a human approach is an arduous work, an automatic method is needed to improve the analysis efficiency. In this paper, a framework is proposed to perform the image semantic segmentation of transmission lines and their accessories to generate the final defect detection results. A segment connection algorithm based on matrix operations is proposed to rapidly connect the segment features of objects. In line accessory detection, a background filter and an artificial contour segment feature generator are constructed to improve the detection performance. In addition, a distance threshold parameter automatic tuning mechanism is presented. Images provided by China Southern Power Grid Company taken by the UAVs are employed to validate the effectiveness of the proposed image semantic segmentation framework.

**INDEX TERMS** Transmission line, unmanned aerial vehicle, semantic segmentation, object detection, image analysis.

## I. INTRODUCTION

Transmission lines serve in hostile environments in which they are exposed to lightning strikes, chemical contaminant corrosion, external destructive forces and so on [1], [2]. After an extended service period, the components of a transmission line will age and fail, which increases its failure probability. Transmission line failures can cause power interruptions and even large-scale grid blackouts, resulting in serious capital losses. To reduce the above risks, it is necessary to inspect the transmission lines and find hazardous defects in a timely manner.

Transmission lines often pass through vast areas, including primitive forests, large rivers, and steep mountains that are difficult to access, and thus their manual inspection has always been a challenging and labor-intensive task in the operation and maintenance of a power grid. Several alternative approaches have been introduced to facilitate inspection, such as robot inspection [3], [4], ultrasonic guided wave inspection [5], and acoustic inspection [6].

As unmanned aerial vehicle (UAV) technology has become mature, deploying UAVs to replace manual labor for inspect-

ing the health of transmission lines has been tried extensively. UAVs equipped with digital cameras can be remotely controlled, allowing them to freely scan the transmission line and wirelessly transmit the captured images/videos [7]. The main advantages of applying UAVs to inspect transmission lines include the following:

1. The personnel injury risk is considerably reduced during the inspection;
2. The inspection efficiency is significantly improved;
3. Transmission line information is recorded as an image/video clip that can be repeatedly presented.

UAVs can improve the efficiency of collecting information on the health of transmission lines; however, an analysis of the collected images is usually performed by a time-consuming manual approach. Thus, it is logical to develop an automatic image semantic segmentation framework to process the images, which lays the foundation for further transmission line detection and explorations of other surface defects of transmission lines.

There have been many meaningful attempts to employ image processing technology in power transmission system inspection. In [8], a sequential local-to-global power line detection algorithm is proposed. Additionally, in [9], an algorithm for power line identification and extraction from

The associate editor coordinating the review of this manuscript and approving it for publication was Canbing Li.

satellite images based on an improved Radon transform is proposed. In [10], the application of big data technology in the fast image recognition of transmission towers based on convolutional neural network (CNN) is researched. Reference [11] presents a vision-based approach for both the identification and the estimation of the relative distance between a UAV and a transmission tower. Taking the spatial correlation between a transmission line and a tower into account, a novel transmission line detection method is proposed in [12]. Additionally, in [13], a new method of feature description for insulator species recognition is proposed. Based on UAV videos, [14] presents a design and implementation of a vision system for the detection of power transmission components, including the transmission line, tower, and insulator.

For the further detection of the defective and damaged components of transmission lines, a digital signal processing (DSP)-based image monitoring method for the high-voltage transmission lines is presented in [15]. In [16], an algorithmic solution for bunch-drop fault detection for both glass and ceramic insulators is proposed based on spatial morphological features. In addition, based on CNNs, intelligent discriminant diagnosis methods for insulators are proposed in [17]–[20].

The most related studies to our investigations are in [21] and [22]. A broken strand detection method that can be practically applied by maintenance robots is presented in [21]. In [22], a multilevel perception-based method of transmission line structure recognition is proposed. However, the images taken by UAVs often have various perspectives and shooting distances with complex backgrounds, leading to unsatisfactory detection performance.

In recent years, image recognition based on CNNs has become a research hotspot [10], [17]–[20], [23]; however, CNN-based methods are not applicable in transmission line inspection tasks due to the following limitations:

1. They require large training samples that are labeled by hand. Such samples are sometimes not available, and the labeling is arduous work.
2. Their training and prediction steps rely heavily on high-performance GPUs and thus are unsuitable for ordinary PCs.
3. They detect objects through specific features, but foreign bodies hanging on transmission lines have various forms. Therefore, CNN is not appropriate for foreign body detection.

Considering the specific semantic segmentation task in transmission line inspection, this paper proposes a systematic framework to automatically analyze images of transmission lines taken by UAVs. The contributions of this paper are as follows:

1. A systematic framework is proposed for semantic segmentation of transmission lines and their accessories based on UAV-taken images. The transmission line components, including the transmission line, transmission tower, and line accessories, are detected simultaneously. To the best of the authors' knowledge, this is

the first integrated framework for image semantic segmentation of transmission lines and their accessories.

2. A novel segment connection algorithm based on matrix operations is established. Numerical results have shown that the algorithm runs hundreds of times faster than the loop-based algorithm, which solves the speed bottleneck in the entire detection framework.
3. Our detection algorithm eliminates the troublesome interference of complex backgrounds by a background filter and compensates for missed object contour segment features by an artificial contour segment feature generator. Both the precision rate (PR) and the recall rate (RR) of the detection performance are improved.
4. Taking the width of the transmission line as a reference, an automatic distance threshold tuning mechanism is proposed. This mechanism makes the detection framework more universally applicable for diverse image resolutions and shooting distances.

The remaining parts of this paper are organized as follows: In section II, the entire identification framework is presented. Section III and Section IV describe the detection approaches and procedures in detail, respectively. A comprehensive real case study is conducted, and the computational results are analyzed in section V. The conclusions of this study are presented in Section VI.

## II. FRAMEWORK

The framework of semantic segmentation of transmission lines and their accessories based on UAV-taken images is composed of four phases: segment feature extraction, regions of interest (ROIs) proposal, accessory detection, and broken strand and foreign body recognition, as presented in Fig. 1.

In phase 1, the segment features of the images are extracted using the line segment detector (LSD) algorithm [24], and then the fragmentary segments are quickly connected into complete contour segments by the proposed matrix-based segment connection algorithm.

In phase 2, through the length limitation, boundary approximation examination, and horizontality examination, the segments of the transmission lines are selected, and the tower zones are identified by analyzing the number and angle distribution of the segments spreading across the image. Afterwards, a mask of the ROIs is generated.

In phase 3, within a specific distance under the transmission line, transmission line accessories such as vibration dampers and spacers are identified based on the chains of  $k$  adjacent contour segments (kASs) [25]. A background filter and an artificial contour segment feature generator are constructed to improve the detection performance. To make the detection framework more robust, an automatic distance threshold tuning mechanism is proposed.

In phase 4, broken strands are detected, followed by foreign body detection. In this phase, the background filter and the automatic distance threshold tuning mechanism are reapplied.

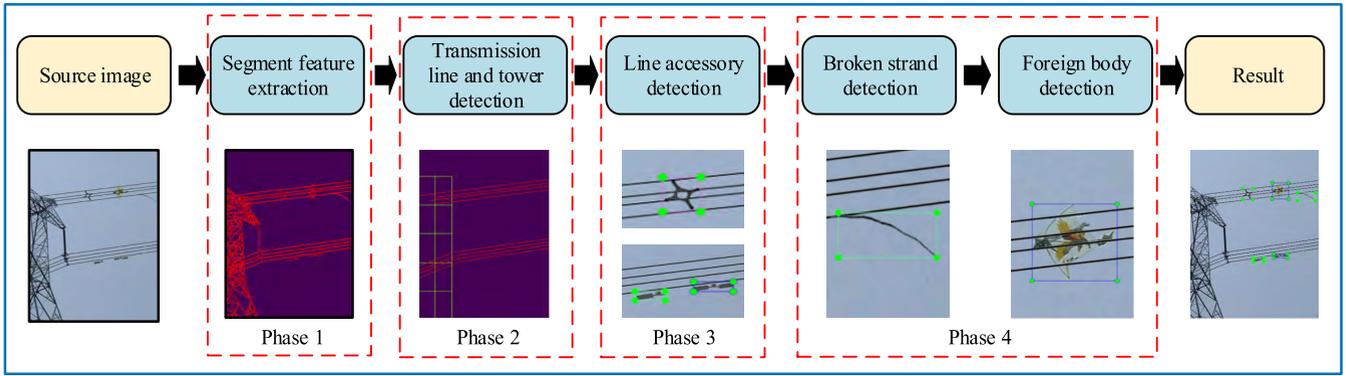


FIGURE 1. Detection framework.

### III. APPROACH

#### A. MATRIX-BASED SEGMENT CONNECTION ALGORITHM

In the object identification field, segment features are often utilized because segment features can be easily represented by the coordinates of their two endpoints while take considerable discriminative information for object detection. The Radon transform (RT), the Hough transform (HT), and the LSD are the most commonly applied algorithms to extract segment features from images [9], [26], [27]. However, the HT and RT are computationally expensive and always result in poor performance with many outliers. Compared to the HT and RT, the LSD has advantages of a higher detection speed, subpixel accuracy, and no parameter tuning. Thus, the LSD is used in this paper. However, an integrated linear contour of an object often produces fragmentary segments because the LSD has a weak tolerance for image noise. To address this problem, segment connection, following segment extraction, is necessary.

The connection of two segments is the basis of segment connection. To determine whether two segments should be connected, the proximity, continuity, and collinearity of the gestalt perception law [28] are employed. As shown in Fig. 2, proximity, continuity, and collinearity are described as (1)-(3).

$$d_{s_1, s_2} < d_{thr1} \tag{1}$$

$$|\theta_1 - \theta_2| < \theta_{thr1} \tag{2}$$

$$d_v < d_{thr2} \tag{3}$$

where  $d_{s_1, s_2}$  is the Euclidean distance of point  $s_1, s_2$  ( $s_1, s_2$  are the two near endpoints of  $L_1, L_2$ );  $d_{thr1}$  is a given distance threshold;  $\theta_1$  and  $\theta_2$  are the angles between  $L_1$  and the horizontal axis,  $L_2$  and the horizontal axis, respectively,

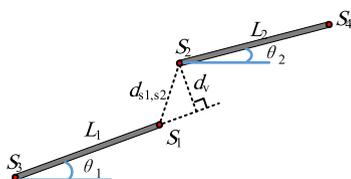


FIGURE 2. Gestalt perception law diagrammatic sketch.

ranging from  $0^\circ$  to  $180^\circ$ ;  $\theta_{thr1}$  is a given angle difference threshold; and  $d_v$  is the distance between  $s_2$  and the foot of the perpendicular on  $L_1$ . If two segments meet all three judgment conditions, then they are connected into a new segment by taking the two farthest endpoints of the two segments as the endpoints of the new segment.

When the segment connection extends to a large number of segments, the procedure becomes complex. A loop-based segment connection algorithm that successively executes two segments connection is possible [29]. However, this loop-based algorithm is excessively redundant, and as a result the connection procedure becomes the performance bottleneck in the whole detection framework. To address this performance bottleneck, a segment connection algorithm of high efficiency based on matrix operations is developed. The main idea of the matrix-based segment connection algorithm is as follows:

*Step 1:* Examine every segment simultaneously with all the others according to gestalt perception law by matrix operations and generate an  $n \times n$  adjacency matrix;

*Step 2:* Extract connected components from adjacency matrix, where each connected component contains segments should be connected;

*Step 3:* Generate new segments using a weighted least square method.

Assume there are  $n$  segments, and if the  $i^{th}$  segment can be represented by coordinates of its two endpoints  $(x_1^{(i)}, y_1^{(i)}, x_2^{(i)}, y_2^{(i)})$ , then matrix  $X_1$  is defined as (4):

$$\begin{bmatrix} x_1^{(1)} & x_1^{(2)} & & x_1^{(n)} \\ x_1^{(1)} & x_1^{(2)} & & x_1^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(n)} \end{bmatrix}_{n \times n} \tag{4}$$

where  $x_1^{(1)}, x_1^{(2)}, \dots, x_1^{(n)}$  stand for the first coordinates of the  $n$  segments, respectively. Matrixes  $Y_1, X_2,$  and  $Y_2$  are defined in the same way as  $X_1$ .  $D_{s1,s2}, D_{thr1}, \Theta, \Theta_{thr1}, D_v,$  and  $D_{thr2}$  are described in (5)-(19), where  $d_{thr1}, \theta_{thr1},$  and  $d_{thr2}$  are derived from (1)-(3). Matrix  $A^T$  means the transpose of matrix  $A$ , and all the matrix operations are

point operations.

$$D_{s1,s2} = \min(D_1, D_2, D_3, D_4) \quad (5)$$

$$D_1 = \sqrt{(X_1 - X_1^T)^2 + (Y_1 - Y_1^T)^2} \quad (6)$$

$$D_2 = \sqrt{(X_1 - X_2^T)^2 + (Y_1 - Y_2^T)^2} \quad (7)$$

$$D_3 = \sqrt{(X_2 - X_1^T)^2 + (Y_2 - Y_1^T)^2} \quad (8)$$

$$D_4 = \sqrt{(X_2 - X_2^T)^2 + (Y_2 - Y_2^T)^2} \quad (9)$$

$$D_{thr1} = [d_{thr1}]_{n \times n} \quad (10)$$

$$\Theta = \arctan\left(\frac{X_2 - X_1}{Y_2 - Y_1}\right) \quad (11)$$

$$\Theta_{thr1} = [\theta_{thr1}]_{n \times n} \quad (12)$$

$$D_v = \min(D_v^{(1)}, D_v^{(2)}, D_v^{(3)}, D_v^{(4)}) \quad (13)$$

$$D_v^{(1)} = \frac{|(X_1 - X_1^T) \cdot (Y_2^T - Y_1^T) - (Y_1 - Y_1^T) \cdot (X_2^T - X_1^T)|}{L^T} \quad (14)$$

$$D_v^{(2)} = \frac{|(X_1 - X_1^T) \cdot (Y_2^T - Y_1^T) - (Y_2 - Y_1^T) \cdot (X_2^T - X_1^T)|}{L^T} \quad (15)$$

$$D_v^{(3)} = \frac{|(X_1^T - X_1) \cdot (Y_2 - Y_1) - (Y_1^T - Y_1) \cdot (X_2 - X_1)|}{L^T} \quad (16)$$

$$D_v^{(4)} = \frac{|(X_1^T - X_1) \cdot (Y_2 - Y_1) - (Y_1^T - Y_1) \cdot (X_2 - X_1)|}{L^T} \quad (17)$$

$$L = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} \quad (18)$$

$$D_{thr2} = [d_{thr2}]_{n \times n} \quad (19)$$

Based on the above definition, (1)-(3) are presented in matrix form in (20)-(22).

$$D_{s1,s2} < D_{thr1} \quad (20)$$

$$|\theta - \theta^T| < \theta_{thr1} \quad (21)$$

$$D_v < D_{thr2} \quad (22)$$

Inequations in (20)-(22) generate three result matrixes, named  $M_a$ ,  $M_b$ , and  $M_c$ , whose elements are bool values, i.e., either *True* or *False*. Finally, the adjacency matrix  $M$  is calculated by (23).

$$M = M_a \cdot M_b \cdot M_c \quad (23)$$

where ‘ $\cdot$ ’ means employing the logical operator *And* in elements of some positions among the matrixes.

An undirected graph is derived from the adjacent matrix  $M$ , and then the connect components are extracted by graph theory. Each of the connected components consists of the segments that should be connected. For each connected component, the weighted least squares method is used to fit the endpoints of the segments, where the weight coefficients linearly increase with the lengths of the segments, and finally obtain a new segment, as presented in Fig. 3.

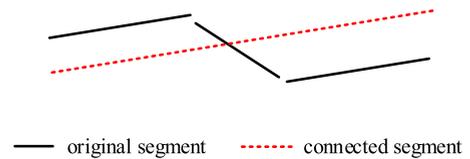


FIGURE 3. Segment connection.

### B. KAS FOR ACCESSORY DETECTION

Many segment features are generated from the image after segment extraction, but the isolated segments take little discriminative information for object detection. To translate the collected information from the segment features as fully as possible, clustering the segments is necessary. There is prior knowledge that a chain of adjacent segments can describe a complex object as long as the chain consists of enough segments. For example, a chain containing three adjacent segments can precisely represent a triangle. Based on prior knowledge, [25] proposes an object recognition method based on a family of scale-invariant local shape features formed by chains of  $k$  adjacent contour segments, called kASs.

As  $k$  increases, the features increase in complexity. On the one hand, the features become increasingly informative, while on the other hand, they gradually become increasingly less repeatable across the different images and object instances. Additionally, the number of non-boundary features also grows with  $k$ , leaving a lower signal-to-noise ratio. Ref. [25] indicates that 2AS is the optimal compromise, while 3AS is slightly inferior. Thus, 2AS and 3AS are introduced to detect line accessories.

Based on the definition of kASs in [25], this paper describes the kASs in a modified way to make the kASs more informative. As shown in Fig. 4, 2AS is described by (24).

$$P_{2as} = (l_1, l_2, \alpha, \theta) \quad (24)$$

where  $l_1$  and  $l_2$  are the lengths of  $L_1$  and  $L_2$ , respectively ( $L_1$  is longer than  $L_2$ );  $\alpha$  is the angle between  $L_1$  and  $L_2$ , computed

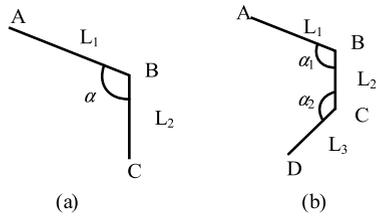


FIGURE 4. Sketch of 2AS and 3AS: (a) 2AS sketch; (b) 3AS sketch.

by (25); and  $\theta$  is the angle between  $L_1$  and the horizontal axis, ranging from  $0^\circ$  to  $90^\circ$ .

$$\alpha = \theta_{\vec{AB}} - \theta_{\vec{BC}} \quad (25)$$

3AS is defined by (26).

$$P_{3as} = (l_1, l_2, l_3, \alpha_1, \alpha_2) \quad (26)$$

where  $l_1$ ,  $l_2$ , and  $l_3$  are the lengths of  $L_1$ ,  $L_2$  and  $L_3$ , respectively ( $L_1$  is longer than  $L_3$ );  $\alpha_1$ , and  $\alpha_2$  are the angles between  $L_1$  and  $L_2$ ,  $L_2$  and  $L_3$ , respectively, calculated in (27)-(28).

$$\alpha = \theta_{\vec{AB}} - \theta_{\vec{BC}} \quad (27)$$

$$\alpha = \theta_{\vec{BC}} - \theta_{\vec{AB}} \quad (28)$$

Referring to [25], the dissimilarity of the two modified 2ASs, and two modified 3ASs are measured by (29) and (30), where  $\omega_\theta$  is the weight coefficient to balance the impact of the angle dissimilarity and length dissimilarity. The parameter  $\theta$  is not considered in the dissimilarity measurement to ensure rotation invariance but will be utilized in other ways. If two kASs pass (31), where  $D_{thr}$  is a given threshold, they are similar.

$$D(a, b) = \omega_\theta |\alpha_a - \alpha_b| + \sum_{i=1}^2 \left| \ln \left( \frac{l_i^a}{l_i^b} \right) \right| \quad (29)$$

$$D(a, b) = \omega_\theta \sum_{i=1}^2 \left| \alpha_i^a - \alpha_i^b \right| + \sum_{i=1}^3 \left| \ln \left( \frac{l_i^a}{l_i^b} \right) \right| \quad (30)$$

$$D(a, b) < D_{thr} \quad (31)$$

### C. ARTIFICIAL CONTOUR SEGMENT FEATURE GENERATOR

In most situations, the UAVs are not straight ahead of the vibration dampers, causing the lateral contours of the vibration dampers to be arc-like rather than straight. As illustrated in Fig. 5(a), the enlarged window presents the lateral contour of the vibration damper and its segment features. In this case, the segment connection algorithm cannot merge the lateral segment features into an integrated one because the large angle difference does not meet the continuity of the gestalt perception law, defined in (2).

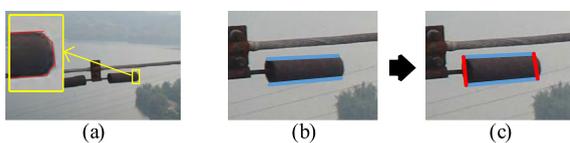


FIGURE 5. Generated artificial segments: (a) The vibration damper's arc-like contours and its segment features (the red segments); (b) Segment features of the vibration damper upper and bottom contours; (c) Original segments and artificial segments.

Note that wherever the UAV collects the images, the upper and bottom side contours of the vibration damper are always straight, and therefore an artificial contour segment feature generator is proposed, if

1. Two segments are parallel and of similar length; and
2. Two segments can vertically project most of themselves to each other.

Then, factitious segments are generated to bridge the two segments. As illustrated in Fig. 5(b), the blue segments are the upper and bottom contour segments of the vibration damper, while in Fig. 5(c), the red segments are the artificial contour segments.

### D. BACKGROUND FILTER

Images are composed of two parts: foreground and background. The main objects to be detected are in the foreground, while the background is the part of the image behind the main objects. If the backgrounds are fields, mountains and so on instead of clear skies, many background segment features will be extracted because the LSD algorithm cannot distinguish foreground and background. Thus, the background seriously influences the detection performance.

A background filter is constructed to handle the above problem. The core component of the background filter is the GrabCut algorithm, which is proposed in [30]. The GrabCut algorithm provides two initialization modes: rectangle initialization and mask initialization. The mask initialization mode is used in this paper because of its advantages in accurateness and flexibility.

A background filter is constructed as follows: First, segment groups are found around the transmission lines. Then, every segment group is encircled by a convex hull. Within the convex hull are pixels that may belong to the foreground; beyond the convex hull are pixels that definitely belong to the background. Given the above convex hull, the GrabCut algorithm is conducted to exclude the background pixels and reserve the real foreground pixels. Lastly, the number of the segment pixels within the convex hull and the reserved segment pixels are counted, and a ruler presented in (32) is proposed to judge whether the object within the convex hull is in foreground.

$$\frac{N_a}{N_b} > \lambda_{thr} \quad (32)$$

where  $N_a$  is the number of reserved segment pixels,  $N_b$  is the number of segment pixels within the convex hull, and  $\lambda_{thr}$  is a given threshold.

Fig. 6 illustrates the sketch of the background filter. There are two groups of segments around the transmission line. In the right group, the GrabCut algorithm reserves most pixels within the convex hull, and the vibration damper passes the background filter. In the left group, GrabCut excludes most pixels within the convex hull, and the gully cannot pass the background filter, and therefore belongs to the background.

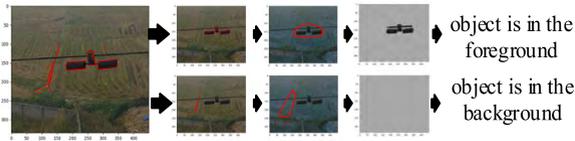


FIGURE 6. Sketch of the background filter.

E. AUTOMATIC DISTANCE THRESHOLD TUNING MECHANISM

In the entire detection framework, many decision criteria refer to distance thresholds. However, a detection framework with constant distance thresholds has little robustness because the UAV images have various resolutions and the shooting distances between the UAVs and transmission lines are different. Note that the actual widths of the transmission lines are invariant, so the width of the transmission lines is chosen as a reference distance, and distance thresholds of the objects detection around a transmission line are scaled according to the transmission line width. The LSD extracts two edge contour segments of the transmission lines, and the distance between the two contour segments is the wanted width.

IV. PROCEDURE

A. TRANSMISSION LINE AND TOWER ZONE DETECTION

Transmission lines and towers are detected after segment feature extraction and segment connection. This paper asserts that among the connected segments, transmission lines are the longest and extend to the boundaries of the images. In addition, transmission lines are nearly horizontal. Based on the three assertions, transmission lines are selected by length limitation, boundary approximation examination, and horizontality examination given in (33)-(35).

$$l_i > l_{thr} \tag{33}$$

$$d_i < d_{thr3} \tag{34}$$

$$\theta_i < \theta_{thr2} \tag{35}$$

where  $l_i$  is the length of connected segment to be judged,  $l_{thr}$  is the given length threshold,  $d_i$  is the minimum Euclidean distance between the segment endpoints and image boundaries,  $d_{thr}$  is the given distance threshold,  $\theta_i$  is the angle between the segment and horizontal axis, and  $\theta_{thr2}$  is a given threshold. (In this paper, the detection programs are accelerated by matrix operations, which are similar to the segment connection algorithm. Due to space limitation, this paper does not elaborate the details and will not mention it again.)

After transmission line detection, the segments are classified into two categories: transmission line segments and non-transmission line segments. All the following object detections are based on non-transmission line segments.

Tower zones are detected by analyzing the number and angle distributions of the segments spreading across the image. The segment features with enough length are selected, and then the selected segments are divided into four classes according to the orientation angles of the segments, as listed

TABLE 1. Grouping criteria for line segments.

class	A	B	C	D
angle(degree)	(-5,5)	(5,85)	(-85,-5)	(-90,-85), (85,90)

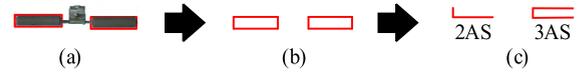


FIGURE 7. Extracting the template kASs of the vibration damper: (a) A real vibration damper; (b) A vibration damper sketch; (c) Template kASs.

in TABLE 1. Afterwards, the images are divided into  $8 \times 6$  zones. Within a zone, if the number of segments belonging to class B and class C are both more than 3, this zone is a tower zone.

After the transmission line and tower zone detection, a ROI mask is proposed to minimize the searching space and improve the detection performance. The regions around the transmission lines are of interest, while the tower zones with a margin are of no interest. The overlapped regions are set as regions of no interest. The mask will be updated along with the entire detection procedure to avoid repetitive work. More importantly, in the following detection of foreign bodies, the ROI mask prevents mistaken recognition.

B. LINE ACCESSORY DETECTION

In this step, line segment features extracted from images are utilized to form kASs. Before vibration damper detection, a vibration damper sketch is needed to extract the template kASs, as demonstrated in Fig. 7, where Fig. 7(a) is a real vibration damper, Fig. 7(b) is the vibration damper sketch, and Fig. 7(c) are the extracted template 2AS and 3AS.

Afterwards, the proposed artificial contour segment feature generator is employed, followed by extracting image 2ASs from the image segment features. As shown in Fig. 8, the image 2AS that meets the following four constraints will be reserved.

- Shape constraint: 2AS, which is similar to the template 2AS according to (29) and (31);
- Angle constraint: the angle  $\theta_v$  of the 2AS, defined in (24), close to the transmission line angle  $\theta_t$ ;
- Position constraint: 2AS, which is under the transmission line, and the ratio of  $d_1$  to  $w$  is within a range;
- Size constraint: the ratio of  $l$  to  $w$  is within a range.

The thresholds of the position and size constraints are available in the vibration damper configuration manual or can be measured from actual images.

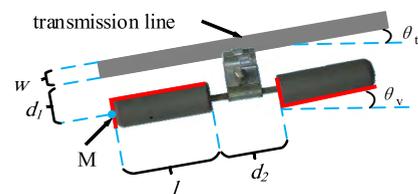


FIGURE 8. Sketch of the vibration damper hanging below the transmission line. M is the middle point of the 2AS short segment.

A 3AS is formed based on two adjacent 2ASs if the two adjacent 2ASs have a shared segment. Only the 3AS that is similar to template 3AS will be reserved, according to (30) and (31). Then, two 3ASs are clustered into pairs if the distance of the two 3ASs  $d_2$ , as shown in Fig. 8, meets (36), where  $d_{thr4}$  is a given threshold. The successfully clustered 3ASs are vibration dampers, while the remaining 3ASs are missing vibration dampers if they pass the background filter.

$$d_2 < d_{thr4} \quad (36)$$

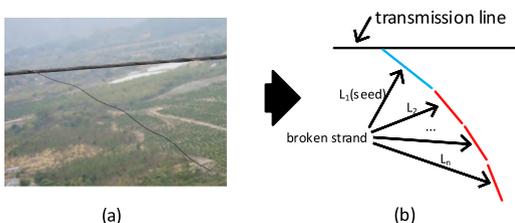
Spacers are detected in a manner similar to the vibration dampers.

### C. BROKEN STRAND DETECTION

Broken lines are the bifurcate line strands that occur on the transmission lines. One of the broken strands' endpoints is close to the transmission line, and the other endpoint is far away. Accordingly, broken strand detection is based on the seed growth method, i.e., segments near the transmission lines are set as seeds, and then the seeds are developed according to (37).

$$d_{si,sj} < d_{thr5} \quad (37)$$

where  $d_{si,sj}$  is the shortest Euclidean distance of the two segments and  $d_{thr5}$  is a given threshold. Fig. 9 illustrates the seed growth method. The background filter is used to examine whether the strand is in the foreground afterwards. Broken strands are detected if seeds can be developed with enough length and pass through the background filter.



**FIGURE 9.** Broken strand detection schematic: (a) Real broken strand; (b) Sketch diagram of the broken strand.

### D. FOREIGN BODY DETECTION

Foreign bodies can be kites, plastic bags, bird nests and so on, which indicates that foreign bodies are diverse without fixed shapes, making foreign body detection a challenge. For the sake of assurance, an object detected near transmission lines is deemed a foreign body. In this case, the ROI mask plays a crucial role because it prevents wrongly recognizing accessories and broken strands as foreign bodies. Segments left in the ROI are first clustered into groups according to the distance limitation defined as (38).

$$d_{si,sj} < d_{thr6} \quad (38)$$

where  $d_{si,sj}$  is the shortest Euclidean distance between the two segment endpoints and  $d_{thr6}$  is a given threshold. Afterwards,

**TABLE 2.** Division of data sets.

	train	val	test	sum
vibration damper	1056	440	1144	2640
missing vibration damper	660	308	308	1276
broken strand	682	154	440	1276
foreign body	836	198	638	1672

the segment groups are examined by the background filter, and the groups in the foreground are foreign bodies.

## V. CASE STUDIES

### A. DATA SETS AND EXPERIMENTAL PLATFORM

We use UAV-taken images whose resolutions are 4000 \* 3000 pixels provided by China southern power grid company to construct our data sets. We label objects using bound boxes with the tag name and then store the label information in.xml files, which are used as ground truth labels. The data sets are divided into training sets, validation sets, and test sets, as listed in TABLE 2. Afterwards, because the data sets are comparatively small, we augment the training sets and validation sets to 3 times the original by horizontal reflection and resizing (scale factor = 0.5).

### B. SEGMENT CONNECTION EFFICIENCY

In this study, we prepare segment groups extracted from different images, whose number of containing segments  $n$  ranges from 3 to 4479. The segment groups are then sent to the loop-based segment connection algorithm proposed in [29] and the matrix-based segment connection algorithms presented in this paper to compare time consumption. Finally, data points are obtained, and every data point contains two subitems,  $t_1$  and  $t_2$ , where  $t_1$  and  $t_2$  are the time consumption in units of seconds of the loop-based algorithm and matrix-based algorithm, respectively.

As presented in Fig. 10, the horizontal axis is the natural logarithm of  $n$ ; the vertical axis is the natural logarithm of  $t_1$ ,  $t_2$ , and  $t_1/t_2$ . When  $n$  reaches 4479, as presented in points A, B, and C, the matrix-based method is approximately 300 times faster than the loop-based algorithm. Since often thousands of segments to be connected are extracted from an image, the matrix-based method reduces the time consumption from hundreds of seconds to several seconds and solves the speed bottleneck.

### C. DETECTION PERFORMANCE

Three metrics, the precision rate (PR), the recall rate (RR), and the harmonic average of the PR and RR F1-measure (F1), calculated by (39)-(41), are applied to assess the performance of the detection framework, where TP is the number of positives, FN is the number of false negatives, and FP is the number of false positives.

$$PR = \frac{TP}{TP + FP} \quad (39)$$

TABLE 3. Parameter analysis.

$c_1$	0.1	<b>0.2</b>	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
F	0.557	<b>0.557</b>	0.557	0.550	0.550	0.548	0.530	0.518	0.518	0.508
$c_2$	1	2	3	4	<b>5</b>	6	7	8	9	10
F	0.476	0.473	0.518	0.536	<b>0.557</b>	0.537	0.522	0.510	0.514	0.481
$c_3$	0.1	0.3	0.5	0.7	0.9	1.1	1.3	<b>1.5</b>	2	2.5
F	0.551	0.572	0.557	0.564	0.572	0.572	0.572	<b>0.572</b>	0.572	0.572
$c_4$	0	0.1	0.2	0.3	0.4	0.5	0.6	<b>0.7</b>	0.8	0.9
F	0.534	0.555	0.555	0.557	0.557	0.572	0.572	<b>0.572</b>	0.572	0.572
$c_5$	5	10	<b>15</b>	20	25	30	35	40	45	50
F	0.551	0.570	<b>0.579</b>	0.572	0.564	0.564	0.557	0.534	0.534	0.534
$c_6$	5	10	15	20	25	30	35	<b>40</b>	45	50
F	0.152	0.258	0.360	0.560	0.579	0.592	0.592	<b>0.592</b>	0.592	0.592
$c_7$	1	2	3	<b>4</b>	5	6	7	8	9	10
F1	0.556	0.581	0.654	<b>0.692</b>	0.679	0.679	0.679	0.679	0.679	0.679
$c_8$	10	20	30	40	<b>50</b>	60	70	80	90	100
F1	0.817	0.891	0.879	0.891	<b>0.894</b>	0.882	0.875	0.884	0.857	0.875

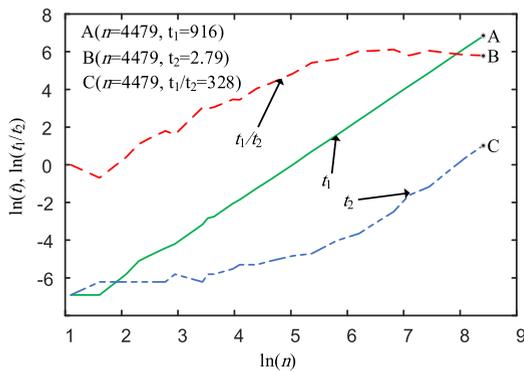


FIGURE 10. Segment connection time  $t$  with  $n$ .

$$RR = \frac{TP}{TP + FN} \tag{40}$$

$$F1 = \frac{2 \cdot PR \cdot RR}{PR + RR} \tag{41}$$

1) PARAMETER ANALYSIS

Our detection framework includes many threshold parameters, of which  $d_{thr1}$ ,  $\theta_{thr1}$ ,  $d_{thr2}$ ,  $\lambda_{thr}$ ,  $l_{thr}$ ,  $d_{thr4}$ ,  $d_{thr5}$ ,  $d_{thr6}$  are the most important. It is vital to analyze the selection of these parameters to obtain the optimum detection performance.

According to our distance threshold tuning mechanism, some of these parameters are related to the width of the transmission line and image resolution. The mapping relations are listed in TABLE 4, where  $w$  is the powerline width;  $a$ ,  $b$  are the width and height of the image resolution, respectively; and  $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$  are the corresponding mapping parameters. Based on TABLE 4, the analysis of  $d_{thr1}$ ,  $\theta_{thr1}$ ,  $d_{thr2}$ ,  $\lambda_{thr}$ ,  $l_{thr}$ ,  $d_{thr4}$ ,  $d_{thr5}$ ,  $d_{thr6}$  is converted to analysis of  $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$ .

We assess  $c_1, c_2, c_3, c_4, c_5, c_6$  by F, defined as (42), where  $F1_v$  and  $F1_m$  are the F1 of vibration damper and missing vibration damper detection, respectively. A small number of

TABLE 4. Parameter mapping relationships.

$d_{thr1} = \max(w * c_1, 5)$	$l_{thr} = \min(a, b)/c_5$
$\theta_{thr1} = c_2$	$d_{thr4} = w * c_6$
$d_{thr2} = \max(w * c_3, 5)$	$d_{thr5} = \max(w, 20) * c_7$
$\lambda_{thr} = c_4$	$d_{thr6} = c_8$

vibration dampers are randomly selected from the validation set. We first assign all parameters default values; next we assign  $c_1$  different values, and select the optimal value according to F; then we reassign the default value of  $c_1$  the optimal value. Default values of  $c_2, c_3, c_4, c_5, c_6$  are reassigned in the same way as  $c_1$  one by one.

$$F = \frac{2 \cdot F1_v \cdot F1_m}{F1_v + F1_m} \tag{42}$$

$c_7$  and  $c_8$  are assessed by F1 of broken strand and foreign body detection, respectively. They are analyzed in the same manner as  $c_1$ . The analysis result of  $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$  are presented in TABLE 3, where the bold items are the optimal values. Based on TABLE 3, the values for  $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$  are set as 0.2, 5, 1.5, 0.7, 15, 40, 4, 50.

2) LINE ACCESSORY DETECTION

In this detection, the background filter and the artificial contour segment feature generator are utilized to improve the detection performance. As a result of the combination, there are four detection methods, A, B, C and D. In addition, we conduct comparisons of our methods and CNN-based methods to validate our methods' competitiveness. In recent years, various CNN-based object detection methods appear, and among them, the Single Shot MultiBox Detector (SSD) [23] is a popular one-stage method, which keeps a good balance between accuracy and speed, benefiting our comparison studies. Though some other CNN-based methods acquire a little higher detection accuracies than SSD, they do not get

**TABLE 5. Confusion matrixes of vibration damper detection.**

	A			B			C		
	V <sub>d</sub>	M <sub>d</sub>	B <sub>d</sub>	V <sub>d</sub>	M <sub>d</sub>	B <sub>d</sub>	V <sub>d</sub>	M <sub>d</sub>	B <sub>d</sub>
V <sub>a</sub>	504	238	402	498	220	426	652	216	276
M <sub>a</sub>	34	155	119	29	149	130	13	185	110
B <sub>a</sub>	138	698	-	75	206	-	234	1003	-

	D			E		
	V <sub>d</sub>	M <sub>d</sub>	B <sub>d</sub>	V <sub>d</sub>	M <sub>d</sub>	B <sub>d</sub>
V <sub>a</sub>	658	203	283	425	21	698
M <sub>a</sub>	13	173	122	22	22	264
B <sub>a</sub>	126	220	-	887	13	-

**TABLE 6. Vibration damper detection performance.**

	vibration damper			missing vibration damper		
	PR (%)	RR (%)	F1	PR (%)	RR (%)	F1
A	74.6	44.1	0.554	14.2	50.3	0.222
B	82.7	43.5	0.570	25.9	48.4	0.337
C	72.5	57.0	0.638	13.2	60.1	0.216
D	82.6	57.5	0.678	29.0	56.2	0.383
E	31.9	37.2	0.343	39.3	7.1	0.121

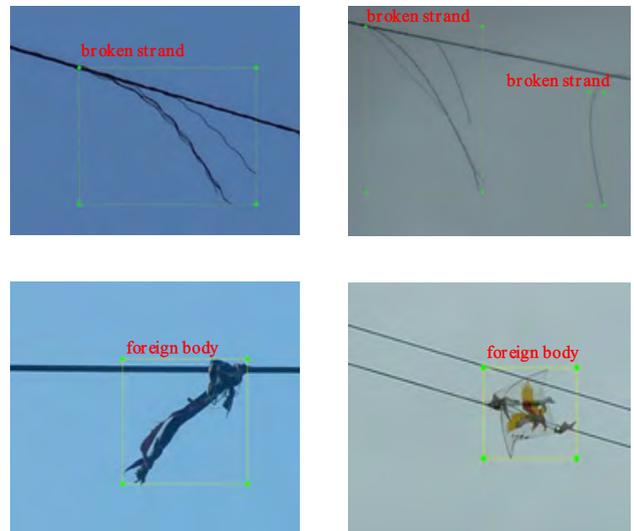
a breakthrough. Thus, the comparison between our methods and SSD is convictive to validate our methods and we refer SSD to here as method E.

- non-improved method;
- improved method with a background filter;
- improved method with an artificial contour segment generator;
- improved method with both the background filter and the artificial contour segment generator.
- Single Shot MultiBox Detector.

The detection confusion matrixes and detection performance of the test set are listed in TABLE 5 and TABLE 6, where V<sub>a</sub>, M<sub>a</sub>, and B<sub>a</sub> stand for the actual vibration damper, the missing vibration damper, and the background, respectively; V<sub>d</sub>, M<sub>d</sub>, and B<sub>d</sub> stand for the detected vibration damper, the missing vibration damper, and the background, respectively; A, B, C, D, and E stand for the detection methods. Some detections are shown in Fig. 11.

As shown in TABLE 6, the comparison of method D and method E demonstrates that our methods surpass SSD overwhelmingly. The F1 of SSD is only 0.343 and 0.121, while F1 of our method is 0.678 and 0.383. In this comparison, the detection performance of SSD is not good because of two reasons: the first reason is that the UAV-taken images are of high resolution (4000 \* 3000), but before CNN feature extraction, the images are resized to standard size (e. g., 300 \* 300), making the details of vibration damper which takes a small part of image seriously lost; the second reason is the relative small training dataset, insufficient for SSD to learn enough knowledge.

The comparison of method D with methods A, B, and C shows that the background filter obviously improves the PR but slightly influences the RR, while the artificial contour generator improves the RR significantly but slightly

**FIGURE 11. Vibration damper detection.****FIGURE 12. Detection of broken strands and foreign bodies.**

influence the PR. This is because the former eliminates most background interference while sometimes mistakenly regards objects in the foreground as background interference; the latter benefits the lateral contour segment feature extraction of the vibration damper while occasionally leading to false detection. The combination of both methods leads to optimal detection performance for the vibration damper, where a high F1 of 0.678 is achieved. However, the PR of missing vibration damper detection is only 29.0%, resulting in a low F1 of 0.383. This is because in the test set, the number of vibration dampers is approximately four times that of the missing vibration dampers, and when a small number of the vibration dampers are mistakenly detected as the missing vibration dampers, a comparatively large FP for the missing vibration dampers is generated. Therefore, the PR of the missing vibration dampers will increase if the ratio of the vibration dampers and missing vibration dampers decreases. Even so, the low PR of the missing vibration damper does not negate the effectiveness of the proposed detection framework because a higher RR of 56.2% is achieved.

### 3) BROKEN STRAND AND FOREIGN BODY DETECTION

The broken strand detection and foreign body detection studies are presented together here. The background filter is employed, and a comparison with SSD is again performed

**TABLE 7. Confusion matrixes of broken strand detection.**

	A		B		C	
	BS <sub>d</sub>	BG <sub>d</sub>	BS <sub>d</sub>	BG <sub>d</sub>	BS <sub>d</sub>	BG <sub>d</sub>
BS <sub>a</sub>	449	145	392	202	155	285
BG <sub>a</sub>	339	-	95	-	80	-

**TABLE 8. Confusion matrixes of foreign body detection.**

	A		B		C	
	Fd	Bd	Fd	Bd	Fd	Bd
F <sub>a</sub>	901	89	883	107	277	361
B <sub>a</sub>	563	-	222	-	22	-

**TABLE 9. Broken strand and foreign body detection performance.**

	broken strand			foreign body		
	PR (%)	RR (%)	F1	PR (%)	RR (%)	F1
A	57.0	75.6	0.650	61.5	91.0	0.734
B	80.5	66.0	0.725	79.9	89.2	0.843
C	66.0	35.2	0.459	92.6	43.4	0.591

in which SSD is denoted method C. In total, there are three detection methods:

- A. non-improved method;
- B. improved method with the background filter.
- C. Single Shot MultiBox Detector.

The detection confusion matrixes and detection performance are listed in TABLE 7, TABLE 8, and TABLE 9, where BS<sub>a</sub>, BG<sub>a</sub>, F<sub>a</sub>, and B<sub>a</sub> represent the actual broken strand, background, foreign body, and background, respectively; and BS<sub>d</sub>, BG<sub>d</sub>, F<sub>d</sub>, and B<sub>d</sub> stand for the detected broken strand, background, foreign body, and background, respectively. Some detection results are presented in Fig. 12.

TABLE 9 validates that the background filter can improve the PR while slightly influence the RR once again. As a result, in broken strand detection, F1 increases from 0.650 to 0.725, and in foreign body detection, F1 improves from 0.734 to 0.843, which are both satisfactory. TABLE 9 also demonstrates that our methods are better than the CNN-based method in broken strand and foreign body detection. In the foreign body detection comparison, another reason for the bad detection performance of SSD is that SSD detects objects through specific features, while foreign bodies have various forms, making SSD not appropriate for foreign body detection.

Finally, we construct an engineering system successfully applying our detection framework to inspect transmission lines, which demonstrates that our framework is feasible for practical application.

## VI. CONCLUSION

Although the deployment of UAVs facilitates the collection of information from transmission lines, an automatic image processing method is needed to efficiently process the images taken by the UAVs. In this paper, an image

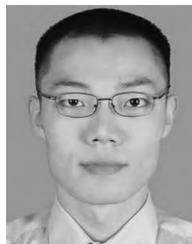
processing framework for semantic segmentation of transmission lines and their accessories is proposed. A matrix-based segment connection algorithm is proposed to accelerate the detection framework, and a background filter and an artificial contour segment feature generator are constructed to improve detection performance. In addition, an automatic distance threshold tuning mechanism is presented to increase the applicability of the framework. Case studies validate the effectiveness of the proposed framework, as the matrix-based segment connection algorithm is hundreds of times faster than the looped-based algorithm, and F1 values for the detection of the vibration damper, missing vibration damper, broken strand, and foreign body are 0.678, 0.383, 0.725, and 0.843, respectively, higher than the corresponding results of SSD detection.

However, this framework is not perfect. The F1 of the missing vibration damper detection is low, and the framework is focused on a specific type of vibration damper and thus is not applicable to other types. Moreover, the detection performance of the tower is insufficient because of the rough detection method.

## REFERENCES

- [1] G. Chen, X. Wang, J. Wang, J. Liu, T. Zhang, and W. Tang, "Damage investigation of the aged aluminium cable steel reinforced (ACSR) conductors in a high-voltage transmission line," *Eng. Failure Anal.*, vol. 19, no. 1, pp. 13–21, Jan. 2012.
- [2] X. Jiang, Y. Xia, J. Hu, Z. Zhang, L. Shu, and C. Sun, "An S-transform and support vector machine (SVM)-based online method for diagnosing broken strands in transmission lines," *Energies*, vol. 4, no. 9, pp. 1278–1300, Sep. 2011.
- [3] O. Menendez, F. A. A. Cheein, M. Perez, and S. Kouro, "Robotics in power systems: Enabling a more reliable and safe grid," *IEEE Ind. Electron. Mag.*, vol. 11, no. 2, pp. 22–34, Jun. 2017.
- [4] N. Pouliot, P.-L. Richard, and S. Montambault, "LineScout technology opens the way to robotic inspection and maintenance of high-voltage power lines," *IEEE Power Energy Technol. Syst. J.*, vol. 2, no. 1, pp. 1–11, Mar. 2015.
- [5] M. K. Yücel, M. Legg, V. Kappatos, and T. H. Gan, "An ultrasonic guided wave approach for the inspection of overhead transmission line cables," *Appl. Acoust.*, vol. 122, pp. 23–34, Jul. 2017.
- [6] K. Park, Y. Motai, and J. R. Yoon, "Acoustic fault detection technique for high-power insulators," *IEEE Trans. Ind. Electron.*, vol. 64, no. 12, pp. 9699–9708, Dec. 2017.
- [7] L. Wang and Z. Zhang, "Automatic detection of wind turbine blade surface cracks based on UAV-taken images," *IEEE Ind. Electron. Mag.*, vol. 64, no. 9, pp. 7293–7303, Sep. 2017.
- [8] B. Song and X. Li, "Power line detection from optical images," *Neurocomputing*, vol. 129, pp. 350–361, Apr. 2014.
- [9] Y. Chen, Y. Li, H. Zhang, L. Tong, Y. Cao, and Z. Xue, "Automatic power line extraction from high resolution remote sensing imagery based on an improved radon transform," *Pattern Recognit.*, vol. 49, pp. 174–186, Jul. 2015.
- [10] Z. Hu, T. He, Y. Zeng, X. Luo, J. Wang, J. Liang, Q. Sun, H. Xu, B. Lin, and S. Huang, "Fast image recognition of transmission tower based on big data," *Protection Control Mod. Power Syst.*, vol. 3, no. 1, pp. 15–25, Dec. 2018.
- [11] O. Araar, N. Aouf, and J. L. V. Dietz, "Power pylon detection and monocular depth estimation from inspection UAVs," *Ind. Robot*, vol. 42, no. 3, pp. 200–213, May 2015.
- [12] J. Zhang, H. Shan, X. Cao, P. Yan, and X. Li, "Pylon line spatial correlation assisted transmission line detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 4, pp. 2890–2905, Oct. 2014.
- [13] W. Yang, H. Y. Yang, and Q. Gao, "Insulator species recognition based on difference features of the color image," *Appl. Mech. Mater.*, vols. 631–632, pp. 441–445, Sep. 2014.

- [14] D. L. Wu, B. F. Li, W. T. Li, Y. Xia, and Y. D. Tang, "A vision-based system for power transmission facilities detection," *Appl. Mech. Mater.*, vols. 423–426, pp. 2547–2554, Sep. 2013.
- [15] C. J. Wang and B. X. Xin, "Abnormal image monitoring for transmission line based on digital image processing," *Adv. Mater. Res.*, vol. 710, pp. 613–616, Jun. 2013.
- [16] 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.
- [17] Y. Liu, S. Pei, W. Fu, K. Zhang, X. Ji, and Z. Yin, "The discrimination method as applied to a deteriorated porcelain insulator used in transmission lines on the basis of a convolution neural network," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 24, no. 6, pp. 3559–3566, Dec. 2017.
- [18] 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)*, Anchorage, AK, USA, Jun. 2018, pp. 1082–1086.
- [19] Z. Zhao, G. Xu, Y. Qi, N. Liu, and T. Zhang, "Multi-patch deep features for power line insulator status classification from aerial images," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Vancouver, BC, Canada, 2016, pp. 3187–3194.
- [20] 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.*, to be published.
- [21] Y. Song, H. Wang, and J. Zhang, "A vision-based broken strand detection method for a power-line maintenance robot," *IEEE Trans. Power Del.*, vol. 29, no. 5, pp. 2154–2161, Oct. 2014.
- [22] Y. Liu, J. Li, W. Xu, and M. Liu, "A method on recognizing transmission line structure based on multi-level perception," in *Proc. 9th Int. Conf. Image Graph. (ICIG)*, Shanghai, China, Sep. 2017, pp. 512–522.
- [23] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, C.-Y. Fu, A. C. Berg, and S. Reed, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.
- [24] R. G. von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "LSD: A fast line segment detector with a false detection control," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 722–732, Apr. 2010.
- [25] V. Ferrari, L. Fevrier, F. Jurie, and C. Schmid, "Groups of adjacent contour segments for object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 1, pp. 36–51, Jan. 2007.
- [26] Z. Xu, B.-S. Shin, and R. Klette, "Accurate and robust line segment extraction using minimum entropy with Hough transform," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 813–822, Mar. 2015.
- [27] U. Budak, U. Halıcı, A. Şengür, M. Karabatak, and Y. Xiao, "Efficient airport detection using line segment detector and Fisher vector representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 8, pp. 1079–1083, Aug. 2016.
- [28] J. H. Elder and R. M. Goldberg, "Ecological statistics of gestalt laws for the perceptual organization of contours," *J. Vis.*, vol. 2, no. 4, pp. 324–353, Feb. 2002.
- [29] Y. Zhang, L. Huo, and H. Li, "Automated recognition of a wall between windows from a single image," *J. Sensors*, vol. 2017, no. 6, May 2017, Art. no. 7051931.
- [30] C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut': Interactive foreground extraction using iterated graph cuts," *ACM Trans Graph.*, vol. 23, no. 3, pp. 309–314, Aug. 2004.



**LONGJUN WANG** received the Ph.D. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2010, he is currently a Lecturer with the School of Electric Power Engineering. His major research interests include power system reliability and planning, and artificial intelligence application.



**ZHAOLONG CHEN** received the B.E. degree in electrical engineering from Fuzhou University, Fuzhou, China, in 2017. He is currently pursuing the master's degree in electrical engineering with the South China University of Technology, Guangzhou, China. His research interest includes digital image processing.



**DONG HUA** received the Ph.D. degree in electrical engineering from the South China University of Technology, Guangzhou, China, in 2012, where she is currently a Lecturer with the School of Electric Power Engineering. Her current research interests include patrol robot and computer vision.



**ZIXUAN ZHENG** is currently pursuing the degree with the International Department, Affiliated High School, South China Normal University. She is currently starting an internship at Guangzhou Intelligence Electrical Technology Company, Ltd. Her research interests include digital image processing and linguistics.

• • •