Novel Vision-Based Abnormal Behavior Localization of Pantograph-Catenary for High-Speed Trains

YIPING LUO [1], QIANRU YANG [1], AND SCARLETT LIU [1,2]

[1]School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China
[2]Department of Engineering, School of Mechanical and Manufacturing Engineering, UNSW Sydney, Sydney, NSW 2052, Australia

Corresponding author: Scarlett Liu (scarlett.liu@csu.edu.cn)

ABSTRACT To ensure the safe operation of high-speed trains, catenary anomaly detection and alerting security have become an urgent problem to solve. In this paper, we propose a novel method for abnormal behavior localization of a pantograph-catenary for high-speed trains. First, a modified faster RCNN is proposed to detect the pantograph faults. By adjusting the parameters of the faster RCNN, the positional accuracy of the candidate box and accuracy of the algorithm are guaranteed. We perform the arc detection after detecting the pantograph head area. The detection accuracy is over 99%. The height of the pantograph center point is also obtained during detection. Then, the actual running mileage of the fault point is calculated. Experiments show that the method proposed in this paper is also applicable to various complex scenes and that this method can determine the fault localization in the shortest time, narrow the maintenance scope, and improve the overhaul efficiency.

INDEX TERMS Deep learning, faster RCNN, fault localization, Pantograph detection.

I. INTRODUCTION

High-speed railways have the advantages of fast speed, less pollution, and large carrying capacity, and have been widely used all over the world [1]. High-speed trains are environmentally friendly compared to other transportation vehicles. However, they must be in continuous contact with a power source. The pantograph-catenary system (PCS) was developed to provide continuous power to high-speed trains. Ensuring continuous contact is a very important in electric high-speed trains [2]. The PCS is a crucial part of high-speed train operation [3]. The train receives power through dynamic sliding contact of the pantograph strip with the catenary. When the train is running and the pantograph strip has poor contact with the contact wire, bow arcing will occur. This arcing generally occurs at higher speeds, increased load, and in cold-weather conditions [4]. The arc results in the following adverse effects: (1) when the bow becomes deformed, a high-amplitude over-voltage is generated, which affects the safe operation of the train; (2) the pantograph strip and contact wire are ablated, shortening their lives and thus leading to

The associate editor coordinating the review of this manuscript and approving it for publication was Yongming Li [ID].

disconnection with the contact line and causing accidents; and (3) electromagnetic radiation and interference are generated. At present, the detection of pantographs in electric high-speed trains throughout the world mainly focuses on the wear of the surface of the skateboard. Most of these techniques use manual visual inspection, which requires the cooperation of multiple staff members. The efficiency and precision are low, the accuracy is poor, and there is a certain safety risk. These methods are also time-consuming and detection can be easily missed due to eye fatigue. Therefore, intelligent anomaly detection for the PCS has become quite important.

Active pantograph systems play an important role in the literature of electrical railway systems. In particular, the use of active pantographs is more suitable at higher speeds. With increasing speed, it becomes difficult to ensure constant contact force. One of the main problems encountered in today's high-speed trains is the loss of contact between the PCS interface due to vibrations in the catenary, which results in reduced current collection. The pantograph was identified by the least-squares method to reduce contact loss [5]. Wu *et al.* [6] used active control to control the system behavior. Anghel *et al.* [7] attempted to control the contact

force by dividing it into two components. Landi *et al.* [8] proposed an infrared-based method to detect the quality of the contact between the pantograph and catenary system. The proposed system used thermal images and applied the canny algorithm to detect edges. Then, the positions of the pantograph and contact point were detected by applying a Hough transform. The conditions of the pantograph-catenary interaction were determined by the temperature of the contact point. Meanwhile, Zhu *et al.* [9] used an edge-detection algorithm to analyze pantograph strips. Any abrasion in the pantograph strip was detected by using five different edge-detection algorithms. An improved pantograph model was proposed by adding wire simulation to apply a constant force to the moving pantograph and overhead line [10]. Furthermore, Allotta *et al.* [11] designed and experimentally verified the complete layout of the PAC system control strategy.

Owing to the rapid development of computer vision and image processing, many researchers are committed to non-contact image detection of the PCS. Karakose *et al.* [12] applied a threshold method to the image matrix consisting of SDRAM and detected arcs. Karakose *et al.* [13] proposed a method based on edge extraction and the Hough transform for obtaining the contact points, and then analyzed the number of contact points in each region of the pantograph to detect PCS faults. In [14] and [15], the mean-shift algorithm and particle swarm algorithm were used separately to track the contact lines. Cho and Ko [16] proposed a dynamic interlacing measurement method based on rotation-invariant feature matching, and achieved good performance under tilting motion. Moreover, Barmada *et al.* [17] extracted abrasion-related features from images of pantograph strips using a discrete wavelet-based method and Hough transform-based method. Additionally, Aydin [18] proposed an automatic detection system for the detecting algorithm. Barmada *et al.* [19] used image processing to detect contact points between the pantograph strip and catenary contact wires to determine if a pantograph arc was present. Ma *et al.* [20] improved the method of detecting wear on the pantograph strip using an image-processing algorithm.

In recent years, the detection of arcs and faults in the PCS has become very important. Signal- and image-processing techniques are used for detection. Some studies in this area have already been mentioned. Hao *et al.* [21] performed a dynamic analysis of the arc in the PCS during the reduction of the pantograph. The PCS was monitored by other authors [22], who solved the problem of adjusting the contact force between the pantograph and the catenary and further proposed a linear time-varying model describing the evolution of contact forces. Meanwhile, Barmada *et al.* [23] proposed a method for detecting arcs in the PCS by using support-vector-based classification. When an arc appeared, they used voltage and current information from the system to obtain a photo type of output. Vázquez *et al.* [24] used a non-contact sensor to monitor the catenary-pantograph effect. The sensor measured the changes in height of the contact wire when the pantograph passed. The deterioration between the

pantograph and catenary was measured. The sensor consisted of a line-scanner camera focusing on an infrared screen. The mathematical model of the sensor system developed in the study and its practical results were given. Ma *et al.* [25] developed a method to calculate the radiation generated by the back of the pantograph on high-speed railways, and combined numerical simulations with laboratory experiments. Capece *et al.* [26] developed an automatic image-based inspection system for locomotive pantographs. Under various environmental conditions, they were capable of capturing and analyzing images at a speed of 300 km/h. They successfully analyzed more than 10,000 pantographs. Wei *et al.* [27] studied high-speed photography and pantograph arcs in a laboratory simulation system, while Tang *et al.* [28] proposed a method to detect visual anomalies in the pantograph head by combining the appearance, scale, and position of the view with the probabilistic Bayesian method. Experimental results showed that the Bayesian detector had better effects than the classical object detector in complex environments. Lederer *et al.* [29] proposed a method for maintaining the contact strip in the pantograph according to the condition of the contact strip. They predicted the wear of the contact strip by monitoring the working distance of the DC component of the locomotive current. Furthermore, Yang and Feng [30] set up a signal- and image-processing-based experiment for pantograph inspection.

The convolutional neural network (CNN) is a feed-forward network that consists of convolutional layers, pooling layers, and fully connected layers. It helps automate image classification [31], object detection [32], and adjustments to scene resolution [33]. In general, the CNN can extract numerous features that are more effective than traditional subjective features. Hao *et al.* [34] proposed an optimized CNN-based image-recognition model and verified its superiority through appropriate experiments. Liu *et al.* [35] used a multi-task multi-view learning method to fuse multiple datasets from different domains, and predicted the water quality of a station. Chen *et al.* [36] proposed a seat-belt detection algorithm for complex road backgrounds. They used multi-scale feature extraction based on deep learning. Meanwhile, Raghavendra *et al.* [37] proposed an 18-layer CNN that effectively extracted powerful features from digital fundus images. These CNN-based methods prove that CNN is a reasonable choice for pantograph arc detection. Karaduman *et al.* [38] proposed a method using deep learning to detect arcs in PCS, and the arc detection was performed using a CNN.

This paper concentrates on computer vision methods to detect faults in the high-speed train pantograph connected to the catenary wire. Object detection based on computer vision is a challenging task, but with important applications. Traditional object detection is based on image color, shape, and texture, with common feature-extraction methods including speeded up robust features (SURF) [39], the deformable part model (DPM) [40], and methods based on the histogram of oriented gradients (HOGs) [41]. Owing to the

complex working environment of a train, different backgrounds, and various interferences, conventional methods cannot achieve good performance. The rapid development of deep learning [42] and CNNs [43] has dramatically improved object-detection performance. Recently, the faster R-CNN [44] was proposed for real-time object detection in pictures. It achieved a mean average precision (mAP) of 73.2% on the PASCAL VOC 2007 and 2012 datasets.

The following is a brief account of the contents of this thesis. Section 2 provides an introduction to the configuration of the experimental environment and the composition of the experimental dataset. Section 3 introduces the modified faster RCNN: by adjusting the parameters and the network structure of the faster RCNN, the positional accuracy of the candidate box and accuracy of the algorithm are guaranteed. The principle of fault localization is also introduced. This section proposes a new fault-localization method based on the scale-invariant feature transform (SIFT) [45]. We use SIFT to perform feature matching between video frames and extract feature-point information, and then calculate the displacement between adjacent frames by feature-point information to further determine the actual geographical localization of the fault point. This is a novel method of abnormal behavior localization. Section 4 presents the results of experiments. The pantograph and arc detection results are given. The actual running displacement of the video frame is also discussed. Section 5 provides the conclusions and contributions of this thesis.

II. ENVIRONMENT CONFIGURATION AND DATASET

The experiment in this thesis configures the py-faster RCNN with Caffe [46] as the deep-learning framework in the Ubuntu 16.04 environment, and the Caffe framework is used to construct the CNNs. CUDA [47], cuDNN [48], Python, OpenCV, and Caffe are used to realize the detection model. The hardware and environment of the experiment are shown in Table 1.

TABLE 1. Computer environment for the experiment.

Configuration	Type
System	Ubuntu16.04LTS
CPU	Intel Xeon Bronze 3104:1.7 Hz, six cores, six threads
GPU	NVIDIA Titan XP:12GB 11400 MHz
Memory	31.1 GB
Disk	SSD 512GB+HDD 4 TB
OS type	64-bit
CUDA	version 9.2.88
cuDNN	version 7.1.4
Python	version 2.7.12
OpenCV	version 3.3.1

We select 1,200 frames from 10 pantograph videos with five different backgrounds and 4,233 frames from 30 arc videos with five different backgrounds for labeling. Their dimensions are 320*240*3. Considering the small number of pantograph datasets, to prevent over-fitting problems the

following operations on the pantograph dataset are performed: transposing, adding obstacles, and adding Gaussian blur. This is shown in Fig. 1.
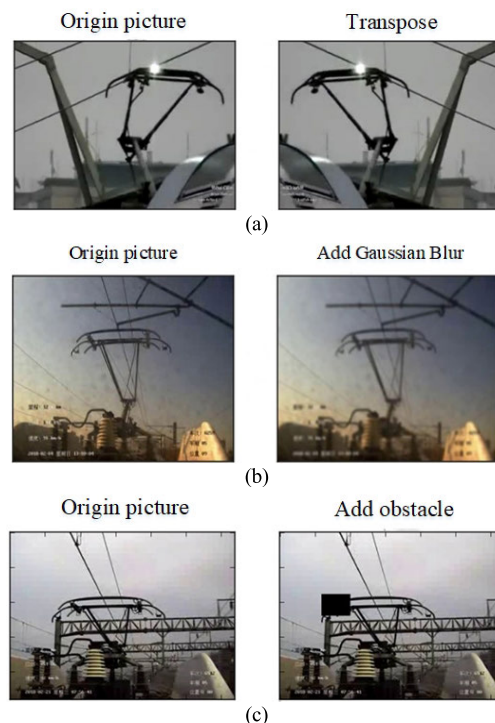

FIGURE 1. Examples of pantograph data enhancement.

The data are calibrated according to the VOC2007 format. The mean of the data must be removed and normalized before using the data. The mean is calculated over the entire dataset. In the process of training deep learning, the dataset is usually divided into a training set, verification set, and test set. The training set learns the sample dataset to match the parameters in order to determine the model parameters. The verification set is used as the evaluation index of the algorithm, and the parameters are adjusted to prevent over-fitting. The test set is used to evaluate the classification effect of the model through the new dataset. In this paper, the trainval set—consisting of the training set and the verification set—accounts for 66% of the entire dataset. The training set and verification set each accounts for 50% of the trainval set. The remaining portion of the test set is used to test the accuracy of the classifier.

III. METHODOLOGY

In this section, the principle of the experiment in this paper is introduced. A modified faster RCNN is proposed for pantograph and arc detection. A novel fault-localization method is proposed. Based on SIFT [45] feature matching, the actual running mileage of the train fault point can be calculated, which saves significant maintenance time and improves maintenance efficiency. Figure 2 shows a flowchart of the method proposed in this paper.
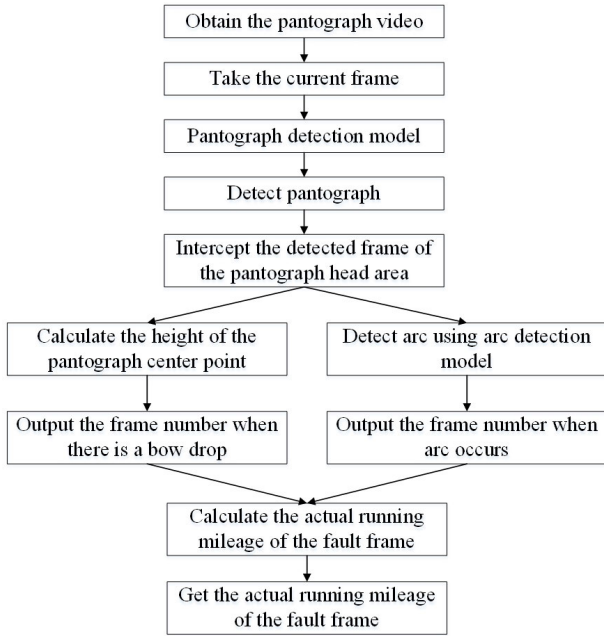
FIGURE 2. Overall view of method proposed in this paper.

A. MODIFIED FASTER RCNN MODEL

The faster RCNN [44] target-detection network is divided into two steps. First, it locates the target, and then it classifies the specific categories of the target. First, the feature-extraction network is used to perform a series of convolution and pooling operations to take the feature map of the image. A region proposal network (RPN) [44] on the feature map locates candidate targets and uses the softmax classifier to determine whether the candidate target belongs to the foreground or background. At the same time, the target candidate range is used to correct the position of the candidate target, and finally the candidate target area is generated. The classification network uses feature maps and candidate regions generated by RPNs to achieve target category detection. Figure 3 illustrates the RPN structure.
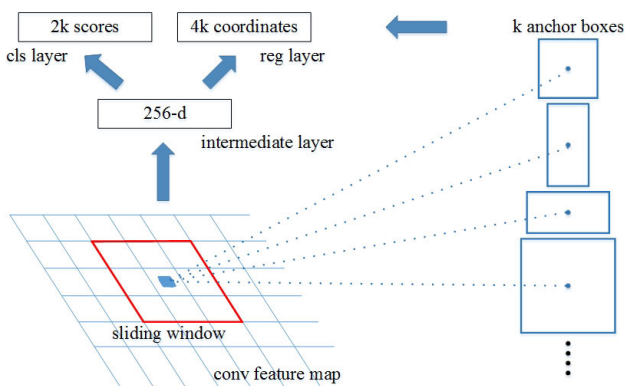

FIGURE 3. RPN structure [44].

When training the RPN, a binary label is assigned to each candidate box for network training, and positive labels are

assigned to the following two cases: (1) the largest candidate box for intersection over-union (IoU) with a real target area box, and (2) the candidate box with an IoU greater than 0.7 for any real target area box. We assign negative labels to candidate frames with IoUs less than 0.3 for all real target candidate frames, and then perform network training and fine tune the parameters. The loss function [51] of the image is defined as

$$L\left(\{p_i\}, \{t_i\}\right) = \frac{1}{N_{cls}} \sum L_{cls}\left(p_i, p_i^*\right)$$
$$+ \lambda \frac{1}{N_{reg}} \sum p_i^* L_{reg}\left(t_i, t_i^*\right). \quad (1)$$

Here, *i* is the index of the ith candidate frame in the small batch processing and p_i the probability that the ith candidate frame is the target. If i is the candidate target, p_i^* is 1; otherwise, it is 0. $t_i = \{t_x, t_y, t_w, t_h\}$ is a vector representing the predicted parameterized candidate frame coordinates. t_i^* is the coordinate vector corresponding to the real target frame. t_i and t_i^* are defined as

$$\begin{cases} t_x = (x - x_a)/w_a, & t_y = (y - y_a)/h_a \\ t_w = \log(w/w_a), & t_h = \log(h/h_a) \\ t_x^* = (x^* - x_a)/w_a, & t_y^* = (y^* - y_a)/h_a \\ t_w^* = \log(w^*/w_a), & t_h^* = \log(h^*/h_a). \end{cases} \quad (2)$$

Here, (x, y) are the coordinates of the center point of the bounding box, (x_a, y_a) are the coordinates of the candidate box, (x^*, y^*) are the coordinates of the bounding box of the real area, and w and h are the width and height of the bounding box, respectively. The purpose of the algorithm is to find a relationship that maps the original box *P* to a regression box that is closer to the real box *G*.

The classification loss function L_{cls} is defined as

$$L_{cls}\left(p_i, p_i^*\right) = -\log\left[p_i^* p_i + (1 - p_i^*)(1 - p_i)\right]. \quad (3)$$

The loss function of the regression L_{reg} is defined as

$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*). \quad (4)$$

Here, *R* is the smooth *L*1 function, which is expressed as

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

The purpose of our work is to achieve real-time monitoring of the pantograph and arc. Since the requirement is real-time detection via video, the processing speed must be addressed. The faster RCNN is modified for this. The ZF network is chosen as the training network. First, local response normalization (LRN) layers are removed. A convolution layer (3*3*63) is added in front of the fc layer of the faster R-CNN to effectively reduce the features, and the dropout is reduced from 0.5 to 0.25.

For the pantograph detection model, the size of the pantograph has remains constant. In high-speed-train pantograph inspection, the monitoring system is located on the train and the camera is opposite the pantograph. The position is fixed

so there is no need to predict regional proposals with different aspect ratios at each sliding window. The aspect ratio of anchors is 1:1 and the anchor's scale remains unchanged.
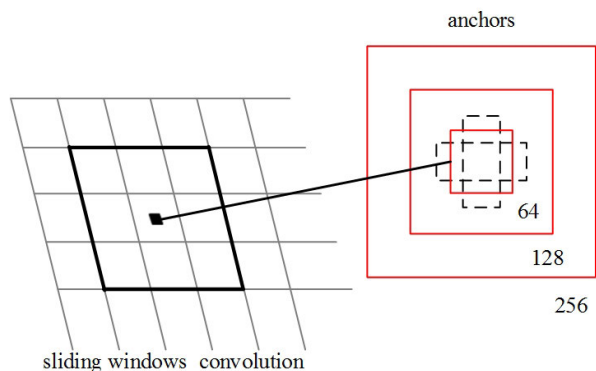

FIGURE 4. Modification of anchor production in RPN of arc detection model.

For the arc detection model, the size of the pantograph arc is generally small, and, thus, in order to make the model more sensitive to small targets, the scales of anchors are changed to 64, 128, and 256 to fit our system. This is shown in Fig. 4. The ratio remains unchanged and nine different candidate regions are generated. In the original faster-RCNN model, the training sample images are all single-scale images. In the case of an arc target with a small footprint in the image, missed detection will occur. To reduce the missed-detection rate during the test, we use multi-scale training. Three input scales (400, 600, and 800) are set. During training, each picture is randomly assigned a scale to input into the network, and therefore the model learning characteristics of the training become wider. Experiments prove that multi-scale training can be used to reduce the missed detection and improve the accuracy of arc detection, so that the trained arc model has certain robustness to the target size.

We train the fast RCNN using 2,000 RPN proposals. The RPN takes a picture of an arbitrary size as input and outputs a certain number of rectangular area candidate frames, in which each area candidate frame corresponds to the probability of presence or absence of a target and the localization information of the target. The traditional RPN passes the 3×3 and 5×5 convolution kernels, generating a total of 256 target feature maps. To reduce the amount of calculations and accelerate the process, we simplify the network structure. As shown in Fig. 5, the modified RPN only passes through a 3×3 volume nucleus to directly produce 256 feature graphs. These 256 feature graphs can be used to obtain 256-dimensional eigenvectors.

In the training process, a four-step alternating training [44] method is used to fine tune the model. First, the network is initialized with ImageNet pre-trained models and fine tuned end-to-end for regional advisory tasks. Then, the RPN generated by suggestion boxes of the first step are used by the faster RCNN to train a separate detection network. This detection network is also initialized by the pre-trained ImageNet model.
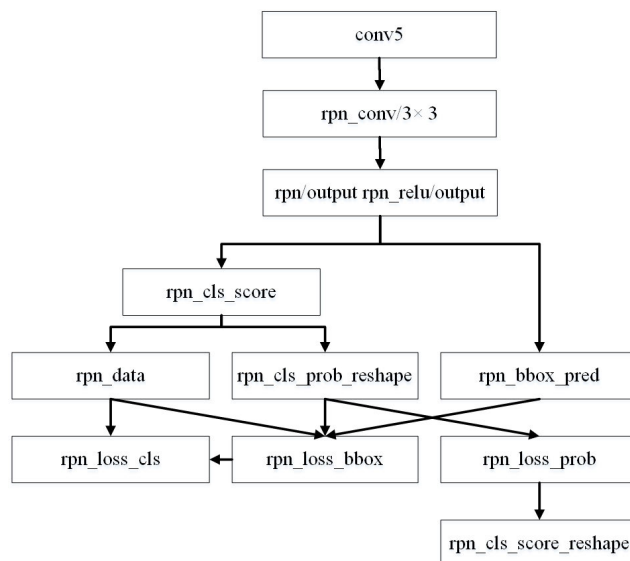

FIGURE 5. Modified RPN.

At this time, the two networks have not shared the convolution layer. Afterward, the detection network is used to initialize RPN training; however, we fix the shared convolution layer and only fine tune the layer unique to the RPN. The two networks now share the convolution layer. Finally, we keep the shared convolutional layer fixed and fine tune the fc layer of the fast R-CNN. In this way, the two networks share the same convolutional layer to form a unified network.

B. FAULT FRAME LOCALIZATION IN REAL WORLD

Considering that there is currently no way to locate the actual geographic localization of a pantograph failure, in this sub-section a novel fault-localization method based on the SIFT method is adopted. The SIFT method is used to perform feature-point matching on the video and extract the coordinates of the matching points. The pantograph camera is fixed on the roof of the train. It is in a relatively static state with the train roof and the pantograph. Figure 6 shows the matching
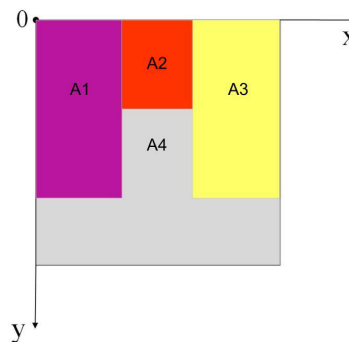

FIGURE 6. Matching areas of pantograph. A1 denotes purple part on left-hand side, A2 red part on upper part, A3 yellow part on right-hand side, and A4 the gray area, which is a relatively stationary area.

areas in our experiment. Points of different matching areas are displayed with corresponding colored lines.

The pantograph video is collected in real time and adjacent frame images are matched in pairs; the coordinates of each pair of matching points in the adjacent frame images are also obtained. By subtracting the coordinates of the paired matching points in the corresponding frame map, the difference between the two matching points at the pixel level is obtained. We divide the sum of the differences between all pairs of matching points at the pixel level by the matching point log, and obtain the mean displacement of all matching points in the two adjacent frame graphs at the pixel level. By analogy, the pixel-level displacement values when the previous frame of all two adjacent frames moving to the next frame are, in turn, obtained.

According to the video-frame diagram of the pantograph fault that has been marked in the foregoing, the pixel-level displacement value S between the first frame and failure frame of the video is obtained. The pixel-level displacement values are added when the previous frame image of all of the two adjacent frame images moves to the subsequent frame. When the first frame moves to the last frame, the pixel-level displacement value $L2$ is obtained. The length of the train's actual running line is $L1$. The actual running mileage of the fault frame from the starting point is expressed as

$$\frac{S}{L2} \times L1. \tag{6}$$

A schematic of video frames matching actual running mileage is shown in Fig. 7, which also shows the matching relationship between video pixel displacement and actual displacement.
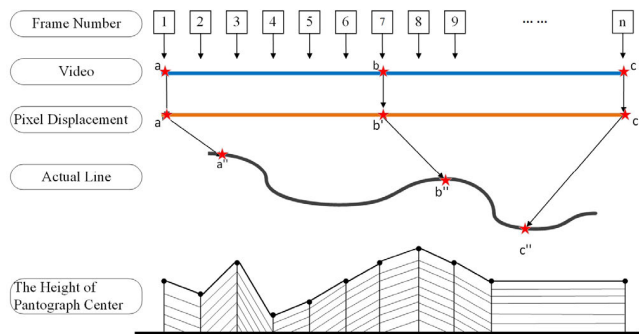

FIGURE 7. Coordinate calibration.

After performing pairwise matching and determining the displacement mean between adjacent frame images, the pixel-level displacement values when the previous frame of all two adjacent frames moved to the next frame are obtained. The displacement of from the first frame image to the second frame image is $D1$, and the movement value from the $N-1$ frame image to the N frame image is D_{N-1}. Assuming that the frame containing the pantograph failure is frame A, the displacement value from the first frame of the video to this frame containing the pantograph failure is

expressed as

$$S = \sum_{i=1}^{A-1} D_i. \tag{7}$$

The total displacement at the pixel level is expressed as

$$L2 = \sum_{i=1}^{N-1} D_i. \tag{8}$$

The length of the train's actual running line is $L1$. The actual geographical position of the fault point corresponding to the A frame map in the actual running line of the high-speed train is expressed as

$$\frac{S}{L2} \times L1 = \frac{\sum_{i=1}^{A-1} D_i}{\sum_{i=1}^{N-1} D_i} \times L1. \tag{9}$$

The proposed method can automatically monitor the state of the pantograph by relying solely on the online camera system installed in front of the pantograph, and can accurately locate the geographical position of the pantograph's failure. Thus, the cost of inspection is greatly reduced.

In summary, the overall framework of our system is shown in detail in Algorithm 1.

IV. EXPERIMENTAL RESULTS

Before training, the parameters of our model are initialized by a model that was trained on the VOC2007 dataset. We comprehensively evaluate the utility on the datasets. It is very important to apply sufficient lifting force to the contact wire while the train is running. The loss of contact will be caused due to vibrations between the pantograph and catenary and oscillations of the contact wire. In this paper, the height of the detected pantograph center point is displayed in real time while taking the pantograph video during high-speed-train operation. When the pantograph is accurately detected, the detection accuracy is displayed above the detection frame, and the change of the center point of the pantograph is displayed in real time in a line graph. We select the pantograph running video in three different running backgrounds to detect the change in pantograph height (Fig. 8), which can be visually observed when the pantograph is raised or lowered. The pantograph's height affects the contact performance between the pantograph and overhead wire. The experimental results of the proposed method show that the detection of the pantograph height is an efficient method for controlling the contact force between the pantograph and overhead line.

Because the work environment of the high-speed train is quite distinct under different scenes, we collected videos under five different scenes for pantograph detection. The frame rate is 25 fps. The scenes are as follows: (1) bright light environment; (2) dim light environment; (3) passing through an access tunnel; (4) crossing a viaduct; and (5) in a maintenance workshop. The detection accuracy in the different scenes is shown in Table 2. We can see that the accuracy is 99% across the various scenarios. Figure 9 illustrates some

Algorithm 1 Overall Framework of Proposed System

Input: Number of frames in video, FN; Fault frame number, $Fault(A)$; Matching area, M;
Matching points of two adjacent frames, M;
Output: Actual running mileage of the fault frame, M;
1.   Load detection video;
2.   **for** $i \leftarrow 1 : FN$ **do**
3.   Read frame x, $y \leftarrow$;
4. $x, y \leftarrow$ the coordinates of all feature points;
5.   **while** $0 \leq x \leq 170$ **and** $0 \leq y \leq 340$ **do**
6. $(x, y) \in A1$;
7.   **end while**
8.   **while** $170 < x \leq 450$ **and** $0 \leq y \leq 60$ **do**
9. $(x, y) \in A2$;
10.   **end while**
11.   **while** $450 < x$ **and** $0 \leq y < 340$ **do**
12. $(x, y) \in A3$;
13.   **end while**
14.   Input the frame of the video to the pantograph model;
15. $x1, y1, x2, y2 \leftarrow$ the coordinates of pantograph area;
16. $Height(i) \leftarrow (y1 + y2)/2$; //Detect the core height of the pantograph
17. $Fault(i) \leftarrow$ the frame number with bowing;
18.   Intercept the frame of the pantograph area;
19.   Input the intercepted picture to the arc model;
20. $Fault(i) \leftarrow$ the frame number with arc;
21.   **end for**
22.   **for** $match(a, b) \in A1\,or\,A2\,or\,A3$ **do**
23. $A \leftarrow Fault(i)$;
24. $S \leftarrow \sum_{i=1}^{A-1} D_i$;
25. $L2 \leftarrow \sum_{i=1}^{N-1} D_i$;
26. $M \leftarrow (S/L2) \times L1$;
27.   **end for**
28.   **Return M**

TABLE 2. Detection accuracy (%) of pantograph.

Scene	Accuracy (%)
Bright light environment	99.4
Dim light environment	99.7
Passing through an access tunnel	99.8
Crossing a viaduct	99.2
In a maintenance workshop	99.5
Average of all scenes	99.8

pantograph-detection examples. The results show that the pantograph detection performs well under different running backgrounds.

To accelerate the detection and reduce detection interference and improve detection accuracy, the model intercepts the video frame with the recorded pantograph position and retains the intercepted partial image of the pantograph as a
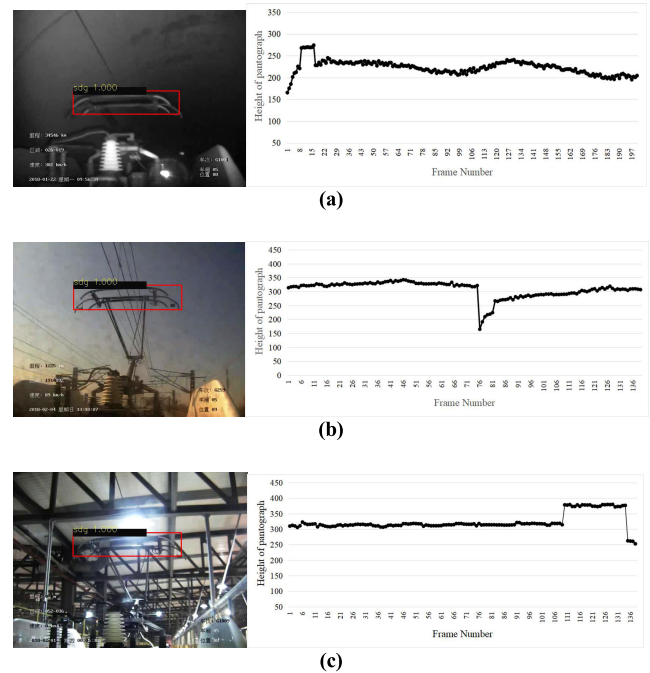

FIGURE 8. Examples of change in pantograph height; (a) variation of pantograph height when passing through a tunnel, (b) change in pantograph height when traversing a viaduct, and (c) variation of pantograph height during workshop maintenance.

TABLE 3. Average precision value for arc test images.

Test Image Number	Average Precision (%)
200	99.4775

detection picture of the arc. Table 3 gives the average precision value of 200 arc test images. The average precision value is found to be 99.4775%, showing how successful the method is in detecting the arc. Figure 10 shows some examples of arc-detection results.

Figure 11 illustrates the pantograph and arc detection process using the modified faster RCNN. We make a comparison between the proposed method and other computer-vision methods as required in Table 4. Compared with the SSD and YOLO models, the processing time and accuracy of SSD are the lowest. Processing time by YOLO is shorter than that by faster RCNN, but the detection accuracy is lower than that of faster RCNN, especially regarding to the detection of small objects (here is the arc). The modified model has higher precision, its performance is relatively better, the processing time is relatively short, and more features of the target generated by the modified model can be extracted.

Two different backgrounds of the pantograph video are selected to perform feature-point matching. Examples of matching are shown in Figure 12.

For the location of the pantograph, we selected a video of a section of the Jiqing high-speed railway, running 30 km from LinZiBeiZhan to QingZhouBeiZhan; it is marked by a red box in Fig. 13.
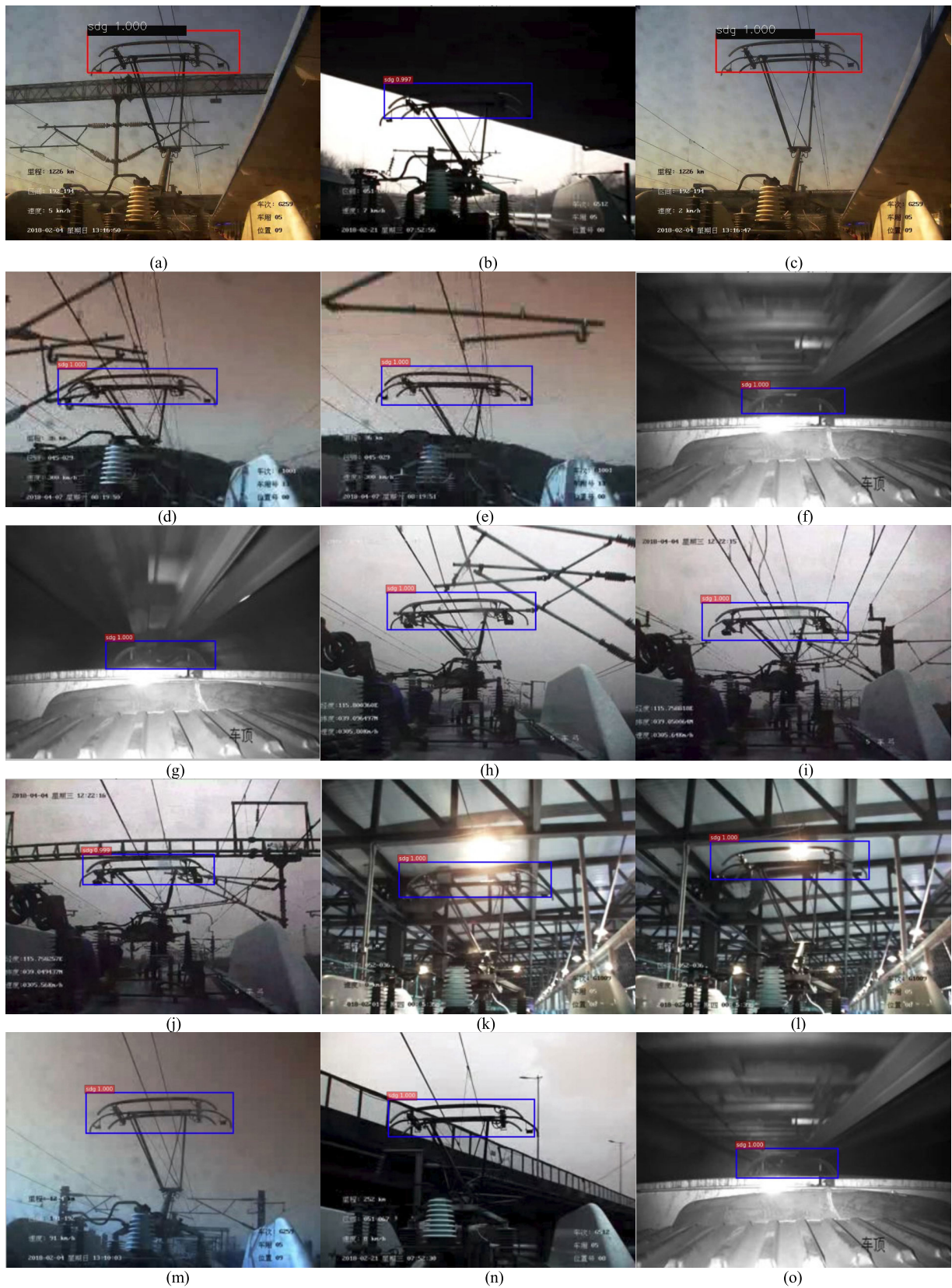
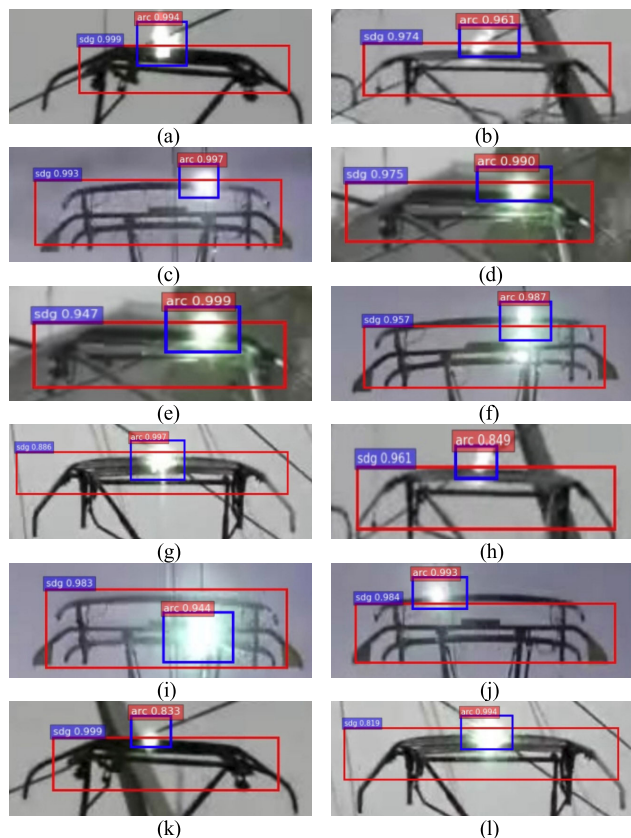**FIGURE 9.** Selected examples of pantograph detection in different backgrounds.
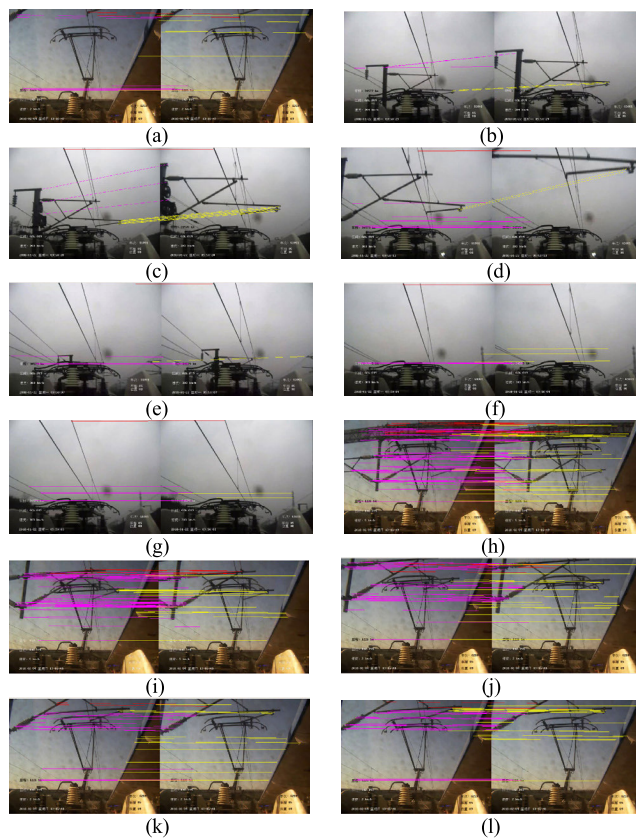
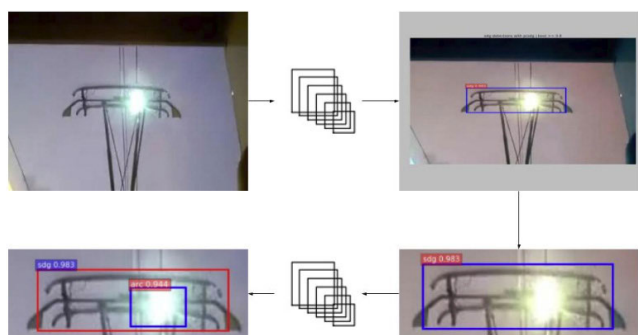**FIGURE 10.** Selected examples of arc detection.

FIGURE 11. Schematic of complete detection process.

TABLE 4. Comparison of different target detection algorithms.

Model	Object	Average Precision(%)	Time (s)
YOLO	Pantograph	93.89	0.019211
	Arc	88.33	0.021149
SSD	Pantograph	90.39	0.061836
	Arc	82.95	0.063187
Faster RCNN	Pantograph	95.31	0.038278
	Arc	94.28	0.038776
Modified Faster RCNN	Pantograph	99.54	0.029776
	Arc	99.48	0.029099

For better display of targeting data, we extracted the actual running line from Fig. 13 and display it in Fig. 14. In Fig. 14, we show the starting and end points. The selected point
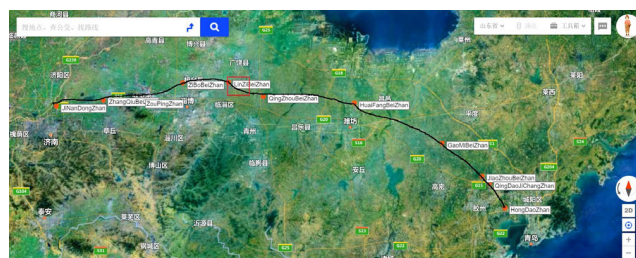

FIGURE 12. Examples of matching in video.

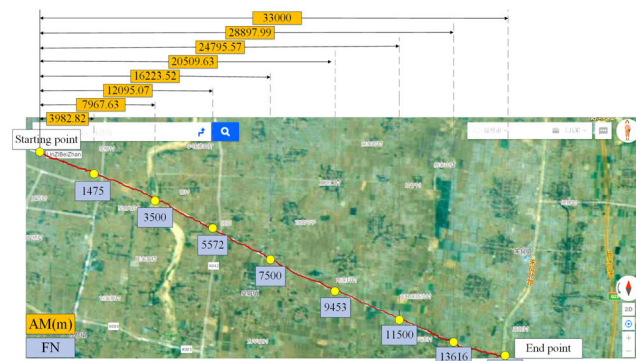**FIGURE 13.** Jinan-Qingdao high-speed railway.

FIGURE 14. Example of localization.

with the actual mileage (AM) and frame number (FN) is determined according to the proposed equation (9).

There are 15,000 frames in the video of the 30-km-long high-speed rail line. The selected example frame data are

TABLE 5. Example of key frame actual running mileage. PD denotes pixel displacement.

FN	PD	AM(m)
1	0.44710307	1.999999995
10	4.188864694	18.73780327
20	11.54531888	51.64499923
40	23.87990986	106.8206031
60	33.23428481	148.6649811
80	44.39667179	198.5970334
100	78.04221166	349.1016577
200	150.0405574	671.1676436
400	280.8718456	1256.407588
600	386.6851346	1729.73598
800	471.7736848	2110.357612
1000	593.0740933	2652.963629
2000	1037.6833	4641.80797
3000	1413.805047	6324.291367
4000	1832.181371	8195.789693
5000	2221.598538	9937.746712
7000	3073.533332	13748.65676
9000	3967.739472	17748.65675
10000	4412.710369	19739.11902
11000	4846.714614	21680.52481
12000	5224.77409	23371.67614
13000	5605.790707	25076.05547
14000	6073.676238	27169.02043
15000	6706.546068	30000

shown in Table 5. This table also shows the pixel-level displacement of the video frames relative to the starting point and displacement relative to the starting point during actual operation. After confirming the actual running mileage of the fault frame, we combine the existing track circuit equipment in the railway field to further confirm the fault location. The track circuit is an important part of a modern railway system [49]. The circuit is used to detect the absence of a train on rails and generate corresponding control signals. A railway track has many block sections, each of which has a separate track circuit. A track-circuit receiver analyzes signals sent from a track-circuit sender and produces a signal to control a relay. When a train is in a section, the relay must be down; otherwise, the relay must be up. We can determine the area where the fault of train pantograph occurred according to the signal from the track circuit and the train running mileage that is calculated. We can then determine the position of the pantograph fault. A significant amount of both maintenance time and cost can be saved by employing the above method.

V. CONCLUSION AND FUTURE WORK
In this paper, we proposed a novel method for abnormal behavior localization of pantograph-catenary for high-speed trains. This study detected the pantograph and the arc, and the localization of the fault of PCS. By first detecting the pantograph and then performing arc detection based on the picture of the pantograph area, the false detection rate of the arc was reduced. Moreover, we could judge the condition of the pantograph through pantograph-height detection. We modified the faster RCNN detection model to match the characteristics of our data. We compared several other

target-detection methods and verified the feasibility and advantages of our method. The accuracy of pantograph detection was 99.58% and that of arc detection was 99.48%. It turns out that the detection accuracy is high and the robustness is strong. In addition, a novel method based on feature matching of the fault localization was proposed, which can locate the fault of PCS in real time, reduce the maintenance scope of the fault, and improve maintenance efficiency. This detection solution can greatly reduce required labor, and it achieves real-time monitoring of high-speed-train operation status. In the future, we plan to detect different faults of PCS and achieve more accurate fault localization, collect different operating scenarios, enrich the database, and provide more complete and accurate identification and localization for the faults of PCS.

REFERENCES
[1] Z. Xu, G. Gao, Z. Yang, W. Wei, and G. Wu, "An Online monitoring device for pantograph catenary arc temperature detect based on atomic emission spectroscopy," in *Proc. IEEE Int. Conf. High Voltage Eng. Appl. (ICHVE)*, Sep. 2018, pp. 1–4.

[2] I. Aydin, E. Karaköse, M. Karaköse, M. T. Gençoğlu, and E. Akin, "A new computer vision approach for active pantograph control," in *Proc. IEEE INISTA*, Jun. 2013, pp. 1–5.

[3] S. Huang, Y. Zhai, M. Zhang, and X. Hou, "Arc detection and recognition in pantograph–catenary system based on convolutional neural network," *Inf. Sci.*, vol. 501, pp. 363–376, Oct. 2019.

[4] S. Midya, D. Bormann, T. Schutte, and R. Thottappillil, "Pantograph arcing in electrified railways—Mechanism and influence of various parameters—Part II: With AC traction power supply," *IEEE Trans. Power Del.*, vol. 24, no. 4, pp. 1940–1950, Oct. 2009.

[5] M. A. Abdullah, Y. Michitsuji, M. Nagai, and G. Venture, "System identification of railway trains pantograph for active pantograph simulation," *J. Syst. Des. Dyn.*, vol. 5, no. 5, pp. 1141–1154, 2011.

[6] Y. Wu, J. H. Zheng, and T. Q. Zheng, "Optimizing active control scheme of high-speed pantograph," in *Proc. IEEE 6th Int. Power Electron. Motion Control Conf.*, May 2009, pp. 2622–2626.

[7] S. Anghel, C. Miklos, M. Topor, D. Demian, and S. Mezinescu, "Pantograph catenary system control using elements of chaos theory," in *Proc. Int. Conf. Pantograph-Catenary Interact. Framework Intell. Control, Conf.*, Jan. 2011, pp. 82–85.

[8] A. Landi, L. Menconi, and L. Sani, "Hough transform and thermo-vision for monitoring pantograph-catenary system," *Proc. Inst. Mech. Eng., F, J. Rail Rapid Transit*, vol. 220, pp. 435–447, Jul. 2006.

[9] X. Zhu, X. Gao, Z. Wang, L. Wang, and K. Yang, "Study on the edge detection and extraction algorithm in the Pantographslipper's abrasion," in *Proc. Int. Conf. Comput. Inf. Sci.*, 2010, pp. 474–477.

[10] A. Balestrino, O. Bruno, A. Landi, and L. Sani, "Innovative solutions for overhead catenary-pantograph system: wire actuated control and observed contact force," *Vehicle Syst. Dyn.* vol. 33, no. 2, pp. 69–89, 2000, doi: 10.1076/0042-3114(200002)33:2:1-1;FT069.

[11] B. Allotta, L. Pugi, and F. Bartolini, "Design and experimental results of an active suspension system for a high-speed pantograph," *IEEE/ASME Trans. Mechatronics*, vol. 13, no. 5, pp. 548–557, Oct. 2008.

[12] M. Karaköse, O. Yaman, I. Aydin, and E. Karaköse, "Real-time condition monitoring approach of pantograph-catenary system using FPGA," in *Proc. IEEE 14th Int. Conf. Ind. Inform. (INDIN)*, Jul. 2016, pp. 481–486.

[13] E. Karaköse, M. T. Gencoglu, M. Karaköse, I. Aydin, and E. Akin, "A new experimental approach using image processing-based tracking for an efficient fault diagnosis in pantograph–catenary systems," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 635–643, Apr. 2017.

[14] I. Aydin, M. Karaköse, and E. Akin, "A robust anomaly detection in pantograph-catenary system based on mean-shift tracking and foreground detection," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2013, pp. 4444–4449.

[15] I. Aydin, M. Karaköse, and E. Akin, "Monitoring of pantograph-catenary interaction by using particle swarm based contact wire tracking," in *Proc. IWSSIP*, 2014, pp. 23–26.

[16] C. J. Cho and H. Ko, "Video-based dynamic stagger measurement of railway overhead power lines using rotation-invariant feature matching," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1294–1304, Jun. 2015.

[17] S. Barmada, M. Tucci, M. Menci, and F. Romano, "Clustering techniques applied to a high-speed train pantograph-catenary subsystem for electric arc detection and classification," *Proc. Inst. Mech. Eng. F, J. Rail Rapid Transit*, vol. 230, no. 1, pp. 85–96, 2016.

[18] I. Aydin, "A new approach based on firefly algorithm for vision-based railway overhead inspection system," *Measurement*, vol. 74, pp. 43–55, Oct. 2015.

[19] S. Barmada, M. Tucci, and F. Romano, "Hierarchical clustering applied to measured data relative to pantograph-catenary systems as a predictive maintenance tool," *Int. J. Railway Technol.* vol. 3, pp. 23–41, Jan. 2014, doi: 10.4203/ijrt.3.4.2.

[20] L. Ma, Z. Wang, X. Gao, L. Wang, and K. Yang, "Edge detection on pantograph slide image," in *Proc. 2nd Int. Congr. Image Signal Process.*, 2009, pp. 1–3.

[21] J. Hao, G. Gao, and G. Wu, "Dynamic analysis of pantograph-catenary arc during the pantograph lowering process," in *Proc. IEEE Int. Conf. High Voltage Eng. Appl. (ICHVE)*, Sep. 2016, pp. 1–4.

[22] C. Lederer, S. Altstadt, S. Andriamonje, J. Andrzejewski, L. Audouin, and M. Barbagallo, "A tracking control for pantograph-catenary system," in *Proc. 54th IEEE Conf. Decis. Control (CDC)*, Sep. 2016, pp. 185–190.

[23] S. Barmada, M. Raugi, M. Tucci, and F. Romano, "Arc detection in pantograph-catenary systems by the use of support vector machines-based classification," *IET Elect. Syst. Transp.*, vol. 4, no. 2, pp. 45–52, Jun. 2014.

[24] C. A. Luna Vázquez, M. M. Quintas, and M. M. Romera, "Non-contact sensor for monitoring catenary-pantograph interaction," in *Proc. IEEE Int. Symp. Ind. Electron.*, Jul. 2010, pp. 482–487.

[25] L. Ma, Y. Wen, A. Marvin, E. Karadimou, R. Armstrong, and H. Cao, "A novel method for calculating the radiated disturbance from pantograph arcing in high-speed railway," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8734–8745, Oct. 2017.

[26] P. Capece, L. Chiesi, A. Pinazzi, M. Sabbatelli, M. Sacchi, L. Venturi, T. Kelsey, and C. Wilson, "PANTOBOT-3D: An automatic vision-based inspection system for locomotive pantographs," in *Proc. 7th IET Conf. Railway Condition Monit. (RCM)*, 2016, pp. 1–5.

[27] W. Wei, J. Wu, G. Gao, Z. Gu, X. Liu, G. Zhu, and G. Wu, "Study on pantograph arcing in a laboratory simulation system by high-speed photography," *IEEE Trans. Plasma Sci.*, vol. 44, no. 10, pp. 2438–2445, Oct. 2016.

[28] P. Tang, W. Jin, and L. Chen, "Visual abnormality detection framework for train-mounted pantograph headline surveillance," in *Proc. 17th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2014, pp. 847–852.

[29] C. Lederer, S. Altstadt, S. Andriamonje, J. Andrzejewski, L. Audouin, and M. Barbagallo, "Condition monitoring of pantograph contact strip," in *Proc. 4th IET Int. Conf. Railway Condition Monit.*, Jun. 2008, pp. 1–6, doi: 10.1049/ic:20080343.

[30] J. Yang and Q.-B. Feng, "A new method for measuring subgrade settlement in high-speed railway by using a linear CCD," *Measurement*, vol. 46, pp. 1751–1756, Jun. 2013.

[31] C. Luo, B. Ni, S. Yan, and M. Wang, "Image classification by selective regularized subspace learning," *IEEE Trans. Multimedia*, vol. 18, no. 1, pp. 40–50, Jan. 2016.

[32] M. Wang, Y. Gao, K. Lu, and Y. Rui, "View-based discriminative probabilistic modeling for 3D object retrieval and recognition," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1395–1407, Apr. 2013.

[33] M. Wang, W. Li, D. Liu, B. Ni, J. Shen, and S. Yan, "Facilitating image search with a scalable and compact semantic mapping," *IEEE Trans. Cybern.*, vol. 45, no. 8, pp. 1561–1574, Aug. 2015.

[34] W. Hao, R. Bie, J. Guo, X. Meng, and S. Wang, "Optimized CNN based image recognition through target region selection," *Optik*, vol. 156, pp. 772–777, Mar. 2018, doi: 10.1016/j.ijleo.2017.11.153.

[35] Y. Liu, Y. Zheng, Y. Liang, S. Liu, and D. S. Rosenblum, "Urban water quality prediction based on multi-task multi-view learning," in *Proc. 25th Int. Joint Conf. Artif. Intell. (IJCAI)*, 2016, pp. 2576–2582.

[36] Y. Chen, T. Gang, H. Ren, X. Lin, and L. Zhang, "Accurate seat belt detection in road surveillance images based on CNN and SVM," *Neurocomputing*, vol. 274, pp. 80–87, Apr. 2017, doi: 10.1016/j.neucom.2016.06.098.

[37] U. Raghavendra, H. Fujita, S. V. Bhandary, A. Gudigar, J. H. Tan, and U. R. Acharya, "Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images," *Inf. Sci.*, vol. 441, pp. 41–49, May 2018, doi: 10.1016/j.ins.2018.01.051.

[38] G. Karaduman, M. Karaköse, and E. Akin, "Deep learning based Arc detection in pantograph-catenary systems," in *Proc. 10th Int. Conf. Electr. Electron. Eng. (ELECO)*, 2017, pp. 904–908.

[39] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," *Comput. Vis. Image Understand.*, vol. 110, pp. 346–359, Jan. 2008.

[40] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminatively trained, multiscale, deformable part model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[41] K. Mizuno, Y. Terachi, K. Takagi, S. Izumi, H. Kawaguchi, and M. Yoshimoto, "Architectural study of HOG feature extraction processor for real-time object detection," in *Proc. IEEE Workshop Signal Process. Syst.*, Oct. 2012, pp. 197–202.

[42] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[43] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 7, 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

[44] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, R. Garnett, Eds. Red Hook, NY, USA: Gurran Associates, 2015, pp. 91–99.

[45] M. Grabner, H. Grabner, and H. Bischof, "Fast approximated sift," in *Computer Vision*, P. J. Narayanan, S. K. Nayar, and H.-Y. Shum, Eds. Berlin, Germany: Springer, 2006, pp. 918–927.

[46] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "CAFFE: Convolutional architecture for fast feature embedding," in *Proc. 22nd ACM Int. Conf. Multimedia*, New York, NY, USA, 2014, pp. 675–678, doi: 10.1145/2647868.2654889.

[47] J. Sanders and E. Kandrot, *CUDA by Example: An Introduction to General-Purpose GPU Programming*, 1st ed. Reading, MA, USA: Addison-Wesley, 2010.

[48] S. Chetlur, C. Woolley, P. Vandermersch, J. Cohen, J. Tran, B. Catanzaro, and E. Shelhamer, "cuDNN: Efficient primitives for deep learning," Oct. 2014, *arXiv:1410.0759*. [Online]. Available: https://arxiv.org/abs/1410.0759

[49] C. Rao, J. Guo, N. Li, Y. Lei, Y. Zhang, Y. Li, and Y. Cao, "Applying combinatorial testing to high-speed railway track circuit receiver," in *Proc. IEEE Int. Conf. Softw. Test., Verification Validation Workshops (ICSTW)*, Mar. 2017, pp. 199–207.

YIPING LUO is currently a Professor with the School of Traffic and Transportation engineering, Central South University. He has published more than 40 academic articles in high-rank academic journals domestically and internationally, which have been indexed and published by the three major academic engines. His main research interests are focus on vision-based intelligent detection and information control for traffic equipment, artificial neural networks and machine learning, advanced design and manufacturing techniques for complex surfaces, rapid prototyping manufacturing and designing, and optimized nesting algorithms and system development. He has completed more than ten national, provincial, and ministerial-level projects as the major research member. He received the second prize of the National Teaching Programs and three Provincial Teaching Program awards.

QIANRU YANG received the bachelor's degree in engineering from the Central South University of Forestry and Technology, in June 2017. She is currently a Graduate Student, majoring in transportation engineering, with Central South University. Her research interests include computer vision and data mining. She is a participant in the National Natural Science Foundation of China (No. 61806222). She has participated in many national competitions and won awards, and has a number of school-level projects.

SCARLETT LIU is currently a Senior Lecturer in vision sensing for field robotics with the School of Traffic and Transportation Engineering, Central South University, China. Her research interests are focused on robotics with experience in robotic vision, sensing, and environment reconstruction. She does theoretical investigation on unsupervised data structure learning from the perspective of vision perception and sensing for autonomous system, which supports decision making and autonomous system control. Her solid publication record and recognition for excellence in presentation skills and teaching confirm her reputation as an Efficient Communicator. Her experience on conducting planning and management of field work accumulated for past five years, proves her capability on efficient filed operations and engagement between academic and industrial parties. Her bilingual skills have been utilized to help improve international recognition of the research team of CV-MLI and supervise international and domestic students. Her clear research plan enhances her ability to set a role model for female engineers/academics in robotics.

● ● ●