IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 4, 2020, accepted November 19, 2020, date of publication December 4, 2020,
date of current version December 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3042535

Online Pantograph-Catenary Contact Point Detection in Complicated Background Based on Multiple Strategies

XUAN YANG[1], NING ZHOU[1], YUEPING LIU[1], WEI QUAN[2], (Member, IEEE), XUEMIN LU[2], (Graduate Student Member, IEEE), AND WEIHUA ZHANG[1]

[1]State Key Laboratory of Traction Power, Southwest Jiaotong University, Chengdu 610031, China
[2]School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China

Corresponding author: Ning Zhou (zhou_ningbb@sina.com)

This work was supported in part by the China National Natural Science Foundation Project under Grant 52072319, in part by the Technology Research and Development Project of China Railway Corporation under Grant P2018J001, and in part by the Scientific Research Project of China Railway Second Academy Engineering Group Company, Ltd., under Grant S2019-32.

ABSTRACT Monitoring train operation status is one of the most important tasks for ensuring rail operation safety. Pantograph and catenary (PAC) are collecting systems of the electric current from traction power supply system, and the stability of the contact between pantograph and catenary guarantees the stable power. However, most existing contact point (CPT) detection methods are always difficult to achieve precise positioning results, especially in complicated background. This article proposes a novel fast and accurate contact point detection method based on multiple strategies, which combine three modules. First, an improved kernelized correlation filter model in real-time tracking module was adopted to track the contact region. Then the pixel-level detection module was used to detect contact point in contact region via the proposed contact point regression residual network (CPRR-Net). Finally, a filter-based optimization module was added to correct the contact position using the Kalman filter. This work additionally employed a new rail dataset PAC-TPL2020 to prove the effectiveness and feasibility of the proposed multiple strategies in real-world scenarios, and the experimental results demonstrated the robustness and high accuracy (97.07% within 3 pixels and 99.97% within 5 pixels) of our model. It is noteworthy that our mothed runs at 65 frames per second for monitoring PAC contact points.

INDEX TERMS Condition monitoring, contact point, deep convolutional neural network, image processing, pantograph and catenary system.

I. INTRODUCTION

In recent years, with the rapid development of high-speed railway systems, condition monitoring and health management became increasingly necessary and important for guaranteeing daily operations. The system supplying and collecting the electric current is the key part for the operation of electrified railways. The stable transmission of electric power guarantees normal railway operations. Consequently, monitoring the operation of the current supplying-collecting system is a key requirement for train operation in railway industry.

Pantograph and catenary (PAC) system [see Fig. 1] is the most used sustained and stable way for modern trains to receive electricity. The operation environment of PAC system is complicated and the parts suffer from mechanical and electrical impact. With the improving speed and long-term operation, there are increasing numbers of accidents occurring in PAC systems, such as arcing, droppers fall-off, abnormal wear of the pantograph strip, even more seriously breakdown of pantograph and contact line fall-off, which cause serious damages to operation companies, even endangering passengers [1]. Therefore, a PAC monitoring system is crucial to ensure the long-term and stable operation of electrified railways.

Contact point (CPT) monitoring is a key part in the whole monitoring system, by which we can calculate the height of operating pantograph and the zig-zag. Moreover, the high-accuracy CPT position can be utilized for the calculation of

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

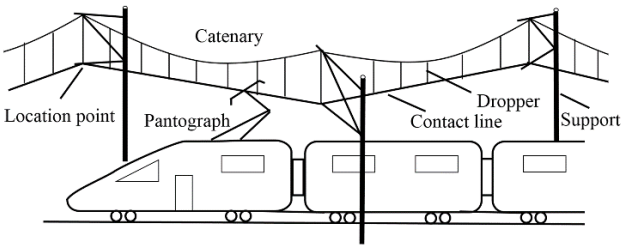


FIGURE 1. The structure of PAC system.

contact force between pantograph and catenary [2]. Hence, CPT detection has become an indispensable part in modern monitoring systems to evaluate the performance of PAC systems. The monitoring of the CPT is the present problematic of current collection in electrical railways.

Traditional CPT detection was based on the detection of pantograph and catenary separately and the CPT was obtained by calculating the intersection point of two lines. The contact line was usually detected in vertical direction and the surface of the pantograph head was detected as the top line by applying the Hough Transform and the Radon Transform [3]–[5]. For the detection of contact line, the mean-shift algorithm [6] and the particle swarm algorithm [7] were applied in other methods and the firefly algorithm [8] was also used for the detection of the pantograph. Template matching was applied to detect parts in PAC system. Zhou *et al.* [9] proposed a dynamic target template matching method to identify catenary suspension movements based on location point feature.

With the development of High-speed Railway, high-speed and complicated background augment the detection requirements and bring more challenges to contact point detection. Traditional methods have a good precision, but they don't perform well if the background is complicated. For example, line detected based method can't detect correctly contact line and pantograph head if there are other lines in the background, especially during the passage of the location point. Template matching are robust to the background but the detection speed and accuracy perform not as well as the line detection. Consequently, it is of great significance to propose a novel method to accurately identify and locate the contact point for further research on PAC system fault diagnosis. In recent years, deep learning-based methods have been widely applied for railway defect detection [10], [11], [12] and greatly improved the accuracy and stability of surveillance methods.

With the development of deep learning, deep convolutional neural networks (DCNNs) play essential roles in image processing and more powerful algorithms have appeared one after another, such as the YOLO series networks [13], [14], [15] and Faster R-CNN [16] for object detection, SiamFC series network [17] for multiple object tracking, which are widely used in the engineering field. In railway PAC system monitoring, DCNNs are also researched and applied for its powerful representation capabilities in computer vision applications, for example, dropper detection[18], strip wear [19], [20] and arc detection [21].

DCNNs achieved performance improvements with development of deep network. Since AlexNet [22] won the LSVRC 2012 classification competition, the most advanced DCNN architecture has grown deeper. AlexNet has only 5 convolutional layers, while the subsequent VGG network [23] and GoogleNet [24] have 19 and 22 layers respectively. However, the deep improvement of the network cannot be achieved through simple stacking of layers. Due to the gradient disappearance problem, deep networks are difficult to train. Because the gradient propagates back to the previous layer, repeated multiplications may make the gradient infinitely small. As a result, as the number of layers in the network becomes higher, its performance tends to saturate, and even begins to decline rapidly. He *et al.* [25] proposed the residual learning framework (ResNets) to ease the training of networks that are substantially deeper than those used previously. ResNets make it possible to train hundreds or even thousands of layers and still show superior performance in this case.

To improve the contact point detection performance, many researchers applied DCNN in detecting the CPT. Zhang *et al.* [26] proposed a method combining deep convolutional network with handcrafted features to detect the CPT. The upper surface of the pantograph strip was detected by an improved DPN and optimized. The edge detection and Hough Transform were used and the widest line was considered the contact line above the pantograph which is not, however, always the case for real application, for example, if the camera is far from the line. Shen *et al.* [27] worked on CPT's real-time tracking in the image sequences with complex background. The method was combined kernelized correlation filters, results refinement and result confidential evolution. But the hypothesis assumes that the CPT's position relative to the region's center is a constant among adjacent image frames, which was inaccurate under certain circumstances.

To solve the above problem and improve the detection speed, we proposed three modules in novel contact point detection method without detecting pantograph head and contact line separately. In the first real-time tracking module, a revised version of KCF algorithm combined with template matching was applied. It combines previous position information and keeps a robust yielding of contact region. Besides it reduces the following computing consumption and improves detection speed. In the second pixel-level detection module, CPT was detected straightly by regression in contact region through a new designed network, called contact point regression residual network (CPRR-Net). In the last filter-based optimization module, a CPT motion theoretical equation was proposed as predictive theoretical value and combined with detected value to located final CPT position in Kalman filtering, which enhances the robustness and accuracy of the model. The combination of the three modules guarantees the robustness of the program against background

interference and also increases the running speed to ensure real-time performance.

In summary, the contributions of this article are as follows:

1) The new proposed method bases on Tracking-Detection-Optimization strategy, which ensures the stability and real-time performance of the algorithm as a whole.

2) We propose a novel contact point regression residual network (CPRR-Net) to yield CPT coordinate in contact region and a CPT motion equation combined with filtering method, which can better locate and correct CPT coordinate under various background conditions. To the best of our knowledge, this is the first work in CPT coordinate regression and in CPT coordinate optimization by Kalman filter.

3) We establish a new dataset PAC-TPL2020, whose conditions mainly consist of arcing, unfavorable weather, light interference, multi-lines interference, etc. The effectiveness of the proposed methods was analyzed thoroughly based on the dataset PAC-TPL2020.

The rest of this article is structured as follows. Section II introduces the structure of the PAC system and presents the current problems encountered in real application scenarios. Section III explains the implantation details of the proposed method for contact point localization and Section IV shows the results from the experiment on real railway lines. Section V provides the conclusion of this article.

II. PAC SYSTEM AND MONITORING CHALLENGES

In this section, we will introduce the pantograph and catenary system, especially the contact parts. Then we introduce the difficulties faced in real application scenarios and give an overall view of the problems that our proposed method aims to solve.

Nowadays, the PAC system is the most adapted ways for trains to receive electricity. The structure is shown in Fig. 1. The catenary system is set along the rail which is composed of messenger, dropper, overhead line, pillar, location point, etc. The pantograph is composed by frame and head. The pantograph head is lifted by the frame and contacts to the overhead contact line. The contact strip on the pantograph head slides along the contact line to get the electricity from the catenary when train goes forward. The contact line is installed as a zig-zag shape along the railway, which aims to reduce wear because the whole contact strip could contact to the overhead wire.

Catenary system adopts double overhead lines in some railway lines. Double overhead lines have their characteristic advantages, for example, augmenting the contact region, which helps reduce contacting wear. A more important reason is that double overhead lines setting has less power loss while transmitting the same amount of power because double overhead lines have less resistance.

The contact points locate at the intersection of the contact lines and the upper surface of the pantograph carbon contact strip. The double contact lines system has two contact points with the pantograph as shown in Fig. 2. However, the contact line itself has a certain width in the collected precise

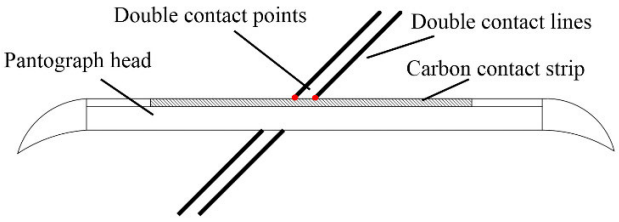


FIGURE 2. Double contact lines model.

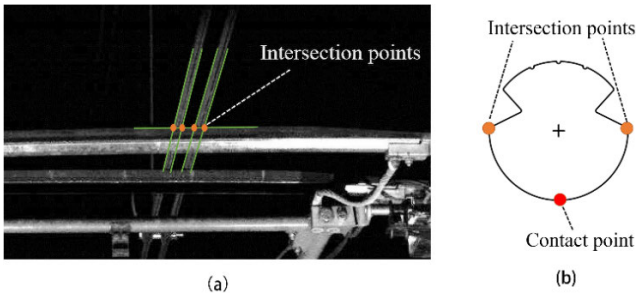


FIGURE 3. Demonstration of contact line structure and key point position. (a) Four intersection points of marginal lines; (b) Cross section of contact line.

image data [see Fig. 3-(a)]. Consequently, there are 2 detected marginal lines for a single contact line and 4 lines for double contact line system based on line detection method. Considering that the bottom of the contact line is arc-shaped [see Fig. 3-(b)], computer vision-based CPT detection hypothesis assumes that contact points' position relative to the middle of two intersection points of marginal lines.

In real application scenarios, the challenge of recognizing PAC contact points in engineering applications lies in the complexity of the background, such as cantilever and bridge frame interference [see Fig. 4-(a)], unfavorable weather conditions [see Fig. 4-(b)], arcing [see Fig. 4-(c)], foreign light interference [see Fig. 4-(d)], etc. In view of the complexity of the PAC system, it is difficult to continuously and stably detect the CPT positions.

In view of complexity of PAC system background, line detection-based method, which is easily interfered by multiple interference, is not feasible for real-world application. Severe weather and light interferences bring huge challenges to the stability and accuracy of other method. Therefore, it is essential and necessary to propose a novel method based on multiple strategies in aim of accurate and robust CPT detection.

III. THE CPT DETECTION METHOD

Towards the accurate and robust CPT detection in complex background, we proposed a progressive detection method based on multiple strategies which include three modules [see Fig. 5]. In the first real-time tracking module, the contact region was focused and tracked by KCF according to initial position which was located automatically by template matching at the beginning of the program. Then a new-designed

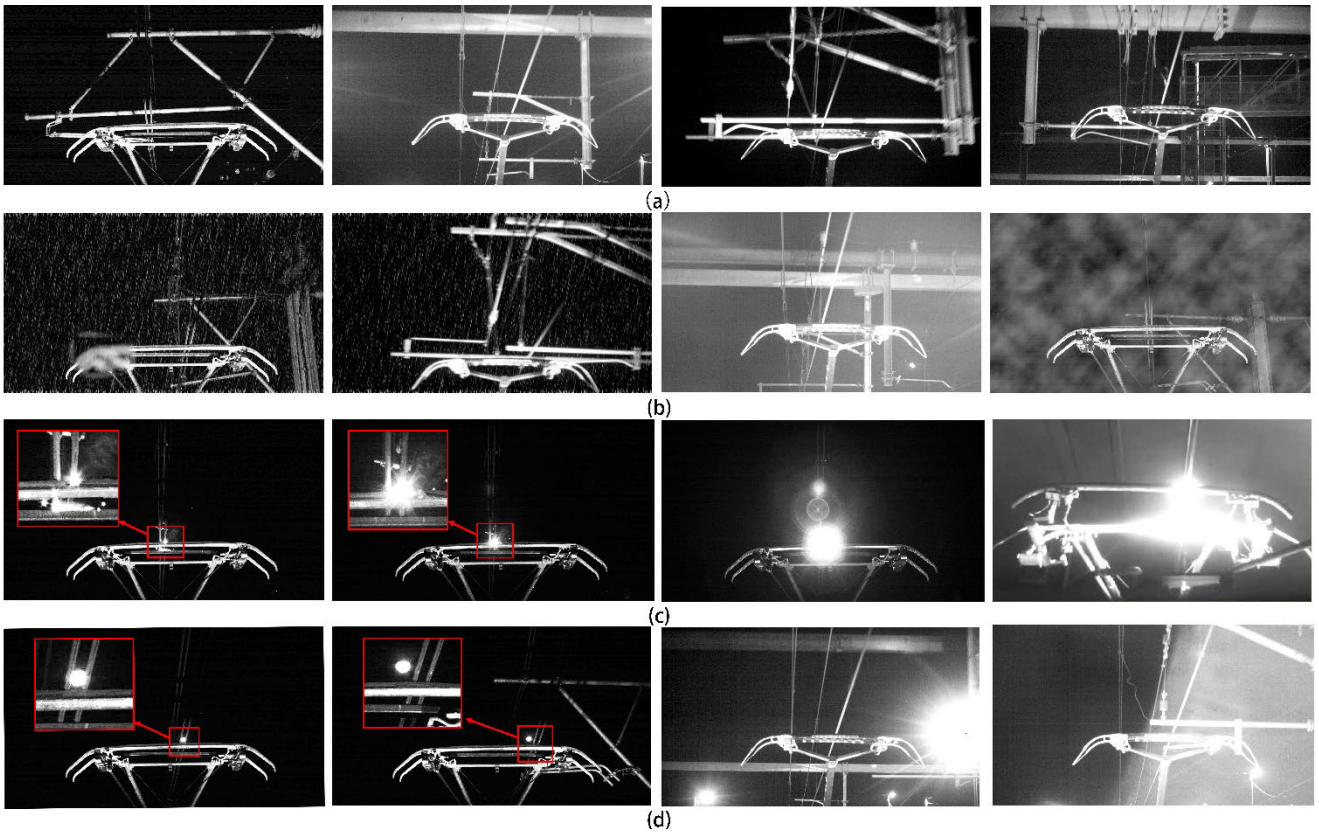


FIGURE 4. Challenges for CPT detection under different backgrounds (Images from dataset PAC-TPL2020). (a) Cantilever and bridge frame interference; (b) Unfavorable weather condition; (c) Arcing; (d) Foreign light interference (The first 2 images are moonlight interference and the last 2 images are road lamp interference).

deep convolutional neutral network, called contact point regression residual network (CPRR-Net) in pixel-level detection module was applied to detect line-intersection points' coordinate in contact region. The outputs of CPRR-Net were the coordinate of line-intersection points and a location point judge index for updating CPT motion equation in following module. In the last filter-based optimization module, Kalman filters corrected contact position by combining the predicted coordinate that was calculated by CPT motion theoretical equation and detected coordinate that was detected by CPRR-Net. The optimization module improved the detection accuracy and eliminated singularities caused by arcing, light interference and other factors.

A. REAL-TIME TRACKING MODULE

In order to realize real-time CPT detection, the tracking method is necessary. Recently, the correlation filter-based single object tracking is one of the most popular and applied methods because of its efficiency and accuracy. However, a single point is impossible to be tracked directly in the complicated railway background, so we first take the contact region with feature of lines intersection as initial targets. On account of the great success achieved by correlation filter-based tracking, KCF [28] was applied in this article to realize the contact region tracking with PAC monitoring images.

In the real-time tracking module, as shown in Fig. 5, a revised version of KCF was adopted to obtain the relevance of the contact region at different displacements by correlation filtering and update the region iterations further through the correlation and send contact region positions frame-by-frame to following steps, then in the next pixel-level detection module, CPTs were located by a method of regression with cropped CPT-focused images, which also helped reduce the amount of data to be processed and improve the detection efficiency.

Before the quick tracking, the initial position containing contact region needs to be transmitted to KCF program in the proposed method. The proposed approach in the tracking module intelligently combined template matching in the spatial domain to locate the contact area, where the characteristics of the intersection between pantograph and catenary was obvious. The matching method distinguished contact lines from messenger lines by its width and gray value because contact lines are thicker and brighter than messenger lines in the acquired images.

Therefore, for the initialization, the filter template which contains contact region moves through the first PAC image and calculates the sum of the gray-scale products of each pixel

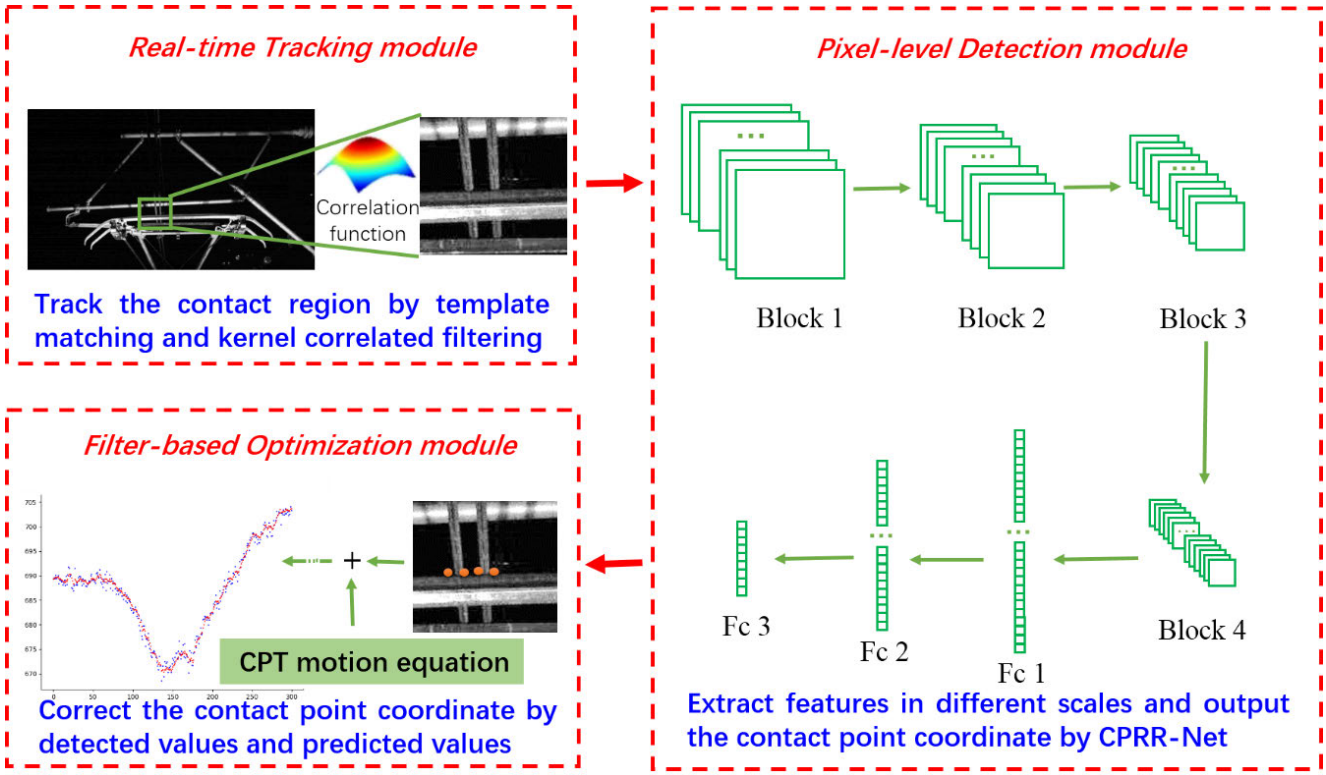


FIGURE 5. Overall flowchart of the proposed methodology for online CPT detection based on multiple strategies.

position, assuming that the size of the image $f(x, y)$ is $M \times N$ and the size of the template sub-image $w(x, y)$ is $J \times K$, then the correlation between f and w:

$$c(x, y) = \sum_{s=0}^{K} \sum_{t=0}^{J} w(s, t) f(x + s, y + t) \tag{1}$$

In the correlation filtering tracking part, the first frame I and the corresponding contact region are provided by matching part. The classifier is trained in Fourier domain using the subsequent training samples, $I_i\,(a, b) \in \{0, \ldots, a-1\} \times \{0, \ldots, b-1\}$, which are captured by cyclic shifts of image I with size $a \times b$ pixels. x is cropped around target center. The expected label y_i(a, b) for each sample x_i follows a Gaussian function, ranging in $0 - 1$. y_i will be 1 for the contact region, and depress to 0 for other regions which shifts away from the contact region. The goal for KCF training is to find a function $f\,(z) = \omega^T z$ that minimizes the squared error between samples x_i and their regression target y_i as follows:

$$\min_{\omega} \sum_{i} (f(x_i) - y_i)^2 + \lambda \|\omega\|^2 \tag{2}$$

where λ is the regularization parameter to control overfitting and ω is the optical value to minimize the equation.

In our proposed method, template matching and KCF cooperated to realize the consecutive tracking of the contact region. As the initialization of KCF, template matching was put into use at the initial segment of each anchor section according to the index p which was the number of spans in an anchor span. When passing overlap spans, template matching was activated with a fixed interval in order to re-locate the new contact line.

With correlation filtering technology in real-time tracking module, the contact region was tracked effectively. The coordinate vectors were transmitted to the clipping section and the origin images were cropped into uniform size, 100 pixel $\times$ 100 pixel in this article. In the end of the tracking module, the cropped images that included contact regions were transmitted to the following detection module.

B. PIXEL-LEVEL DETECTION MODULE

Traditional contact point detection method firstly located the surface of pantograph strip and the contact line separately and then extracted their intersection point as contact point, which process was complicated and computational. To simplify the method and improve real-time performance, the contact point detection was a pixel-pixel regression process in this model. We introduced a novel contact point regression residual network (CPRR-Net), where contact point coordinate was detected directly.

1) LOSS FUNCTION

For the input image I, there were 8 outputted numbers $(\bar{x}_1, \bar{y}_1), \ldots\ldots, (\bar{x}_4, \bar{y}_4)$, representing x and y coordinate values of 4 intersection points and 1 judge index $\bar{u}$ indicating the

detection of location point. The loss function according to the distance between detected values and labels was:

$$Loss = \frac{1}{n}\sum_{i=1}^{n}(x-\overline{x_i})^2 + \frac{1}{n}\lambda\sum_{j}^{n}(y-\overline{y_j})^2 + \gamma(u-\overline{u})^2 \quad (3)$$

where λ is direction-oriented penalty weight whose aim was to ameliorate the detection accuracy in different direction considering that 4 intersection points are distributed in the same straight carbon contact strip line. γ is the weight to improve the accuracy of location point detection. The detection of the location point based on the appearance of registration arm and location clamp in the image.

The final loss function is obtained by the sum of mean square error and weights in the last three full connection layers. During the training validation, accuracy for each epoch was computed by averaging the prediction of batch-size image samples. The gradient descent optimization algorithms were based on Adam [29], an adaptive learning rate method as back-propagation, which has good performance in finding a minimum loss. In the training process of this article, the combination of Adam and loss function realized the fast and reliable regression of CPT coordinate.

2) ARCHITECTURE

The state-of-the-art deep neural network ResNets [25] designed residual learning framework to ease the training of networks and achieved good performance in the areas of image classification and object detection. By inserting shortcut connections which skip one or more layers directly, ResNets realized different deeper layers of network. For example, ResNet-34 consists of 34 convolutional layers which are structured in four composite blocks and each block has 3, 4, 6 and 3 sub-blocks, respectively. There are also global average pooling layers in the network and a fully connected layer in the end of the network. Through such a structured design, Resnets has realized the construction of a deeper network and achieved favorable training results.

To better adapt PAC images and output the CPT coordinate in detection module, we re-architected the ResNets and proposed a novel CPRR-Net, as shown in Fig. 5. From a general view, images in CPRR-Net were fed into 4 blocks in sequence to extract features in different level and followed by 3 fully-connected layer for regression. The new architecture of CPRR-Net takes advantage of the high-level features and also intersection details for detection, and it keeps high accuracy while locating the intersection point between the pantograph upper surface and the contact line.

The implementation of the CPRR-Net was shown in TABLE 1. In details, since the contact region was focused and the intersection feature was enhanced in the last real-time tracking module, the global average pooling layer was replaced by the convolutional layer to remove noises and smooth input images at the beginning of the CPRR-Net. Then the input layer was set to 32 channels instead of the original 3 channels in order to extract characteristics in different angles. Furthermore, the kernel size 1×1 and 3×3 filters were used in most convolutional layers. Each building block in the network was connected to a convolutional layer, followed by a batch normalization (BN) layer, which helped the network training and control gradient explosion or disappear.

TABLE 1. The architecture of CPT regression network.

	Layer	Kernel size	Stride	Output
Pre	Conv0_1	5 ×5	1	100 × 100 × 32
Block1	BN-layer	-	-	100 × 100 × 32
	Conv1_1	1 × 1	1	100 × 100 × 16
	BN-layer	-	-	100 × 100 × 16
	Conv1_2	3 × 3	2	50 × 50 × 16
	BN-layer	-	-	50 × 50 × 16
	Conv1_3	1 × 1	1	50 × 50 × 64
Block2	BN-layer	-	-	50 × 50 × 64
	Conv2_1	1 × 1	1	50 × 50 × 32
	BN-layer	-	-	50 × 50 × 32
	Conv2_2	3 × 3	2	25 × 25 × 32
	BN-layer	-	-	25 × 25 × 32
	Conv2_3	1 × 1	1	25 × 25 × 128
Block3	BN-layer			25 × 25 × 128
	Conv3_1	1 × 1	1	25 × 25 × 64
	BN-layer	-	-	25 × 25 × 64
	Conv3_2	3 × 3	2	13 × 13 × 64
	BN-layer	-	-	13 × 13 × 64
	Conv3_3	1 × 1	1	13 × 13 × 256
Block4	BN-layer			13 × 13 × 256
	Conv4_1	1 × 1	1	13 × 13 × 128
	BN-layer	-	-	13 × 13 × 128
	Conv4_2	3 × 3	2	7 × 7 × 128
	BN-layer	-	-	7 × 7 × 128
	Conv4_3	1 × 1	1	7 × 7 × 512
FC	BN-layer	-	-	7 × 7 × 512
	Fc_1	-	-	512
	Fc_2	-	-	128
	Fc_3	-	-	9

At the end of the CPRR-Net, three full connection layers were designed to output 9 values. The sliding decreasing convolution kernel size was used to convolve the fused feature map in each composite block, reducing the number of channels to 9, and a convolution layer was applied to transform the feature maps into vector size. Then we got the final output of the CPRR-Net. The 9 values represent the x and y coordinate values of 4 intersection points and 1 judge index, which indicates the detection of location point and will be used for updating parameters in the following process.

C. FILTER-BASED OPTIMIZATION MODULE

In the results of CPT coordinate from the previous work, Gaussian noise was inevitable during the processing of both sampling and detection. Therefore, the Kalman filter which combined the predicted results and detected results for

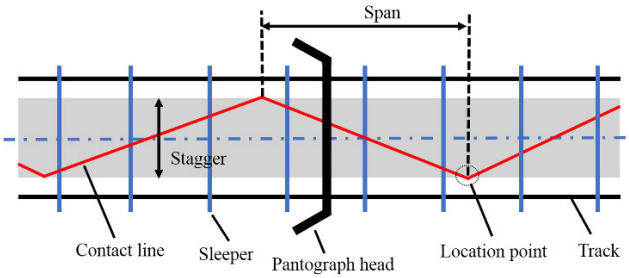


FIGURE 6. Distribution of contact line.

prediction and correction is an effective technique to realize more accurate results. Especially when there were strong arc occurring, the image information was lost and Kalman filter could predict the position in the following frames.

In electrified railways, the contact line is installed as a zig-zag shape along the railway [see Fig. 6] and the contact strip on the pantograph head slides along the contact line when train moves along the railway direction (longitudinal). The lateral movement of the CPT is always kept within the range of the stagger value and CPT trajectory is approximate to a straight line within a span. Therefore, to express the new state of the CPT positions, we proposed CPT state model for prediction step as follows and x is state vector representing position and velocity.

$$\hat{x}_t^- = F_t \hat{x}_{t-1}^- + B_t u_t \tag{4}$$

where $\hat{x}_t^-$ are predictions of the position at time t. F_t is the state transition matrix which defines the relation between the state vectors at time t and $t-1$. B_t is the control matrix and u_t is control vector. The following equation represented the transmission relationship of uncertainty at different moments.

$$P_t^- = FP_{t-1}F^T + Q \tag{5}$$

where P_t is system noise covariance matrix and Q is the covariance representing the untracked influences as noise from the environment.

Through CPRR-Net in pixel-level detection module, we obtained the observed position value z_t and we combined these observed values with the state vector as follow.

$$\hat{x}_t = \hat{x}_t^- + K_t(z_t - H\hat{x}_t^-) \tag{6}$$

$(z_t - H\hat{x}_t^-)$ represented the residual between our actual observations and predicted results. K_t is Kalman coefficient, which is given as follows:

$$K_t = P_t^- H^T (HP_t^- H^T + R)^{-1} \tag{7}$$

In this model, the position of the contact point was only observed, and K already contained the information of the covariance matrix P, so we used the correlation between the two dimensions of position and velocity to infer speed information from the position information. In this way, two dimensions of the state quantity x were modified simultaneously. Therefore, the acceleration u_t was considered as zero

Algorithm 1 Kalman Filter

```
Initialization:
    Get the first two detected point positions from CPRR-Net: (x1, y1), (x2, y2)
    Determine the initial speed of CPT: v0^x, v0^y
    Initialize state vector, covariance matrix, control vector, control matrix
While: ~done
    Get the judge index u from CPRR-Net
    If w> 0.5
        Span count number: m = m + 1
        If m > threshold T   Break
        Else   Re-initialization
    Else
        Predict the next position: x_predict
        Transition of noises: p_predict
        Calculate Kalman coefficient: kalman
        Update contact point position: x_mat
        Update the noise covariance matrix: p_mat
    Show the contact points in the image
Save the vectors as CPT positions
```

in contact point state model. Finally, the noise covariance was updated for the next iteration, which can be expressed as:

$$P_t = (I - K_t H)P_t^- \tag{8}$$

Algorithm 1 introduced the procedures of the combination of predicted results and detected results by Kalman filter to realize the CPT position prediction and correction. For initialization, initial speed was computed by the displacement of CPT point in the first two consecutive frames with sampling interval Δt. The judge index u, which varied in range of 0 and 1, is the ninth output number from the CPRR-Net. The value of w was greater than 0.5 indicating the detection of the location point, and less than 0.5 for other situations which represented contact points within one same span.

With Kalman filter program, gaussian noises in detection results were suppressed and the absent detection points in case of arcing were predicted, which assured the whole program in a state of robustness.

IV. EXPERIMENT

A. EXPERIMENTAL ENVIRONMENT AND DATASET

The CPT position measurement helps monitor operating condition of PAC system and ulteriorly calculate other key parameters. In this article, a novel CPT detection method is proposed and tested based on our newly created dataset, called PAC-TPL2020, which was collected from real operation of railway lines by States Key Laboratory of Traction Power (TPL). The high-resolution image acquisition module was installed on the roof of the train [see Fig. 7]. The visual field of camera focuses on the operating area of the PAC system, and as the train runs, real-world images collected without a break. The resolution of the image is 1920 pixel $\times$1080 pixel and the sampling frequency is 50fps.

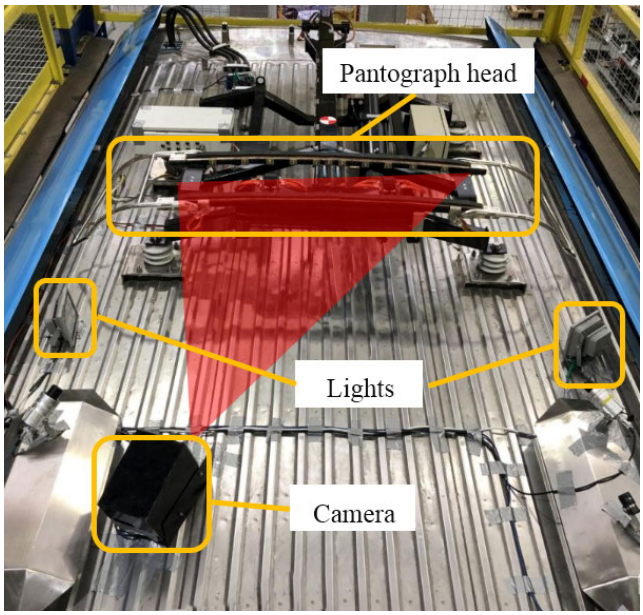


FIGURE 7. The image acquisition module.

In order to detect CPT in various background, we established a new dataset PAC-TPL2020, whose images were acquired in real-world scenarios and calibrated manually with CPT positions. The dataset was not limiting to a specific railway line but instead covered a wide selection of interference in diverse working conditions. For example, PAC-TPL2020 contains different complicated backgrounds, consists of mainly different interference [see in Fig. 4], arcing, unfavorable weather, foreign light interference, cantilever and bridge frame interference, multi-lines interference, etc. Eventually, more than 5000 original PAC system images were calibrated manually in precise level of pixels and the CPT coordinate was enregistered as ground truth. TABLE 2 presented the statistics of different interference. The total number of images was expanded to 5 times in PAC-TPL2020 by data strength technology, for example, adding noises and changing background.

TABLE 2. Statistics of Pac-tpl2020.

Case	Cantilever and bridge frame interference	Severe weather	Arcing	Foreign light interference	Other
Num	799	1311	126	1476	2288

PAC-TPL2020 was separated into 3 parts, training dataset, testing dataset and validation dataset with a ratio of 6:3:1. The diversity of the dataset enabled model architectures to focus on the line-intersection feature between pantograph and catenary in different contexts and the large scale of this dataset enabled the development of deep residual network for CPT regression. All the results in this article were evaluated on this dataset and the richness of data ensure the portability for other lines.

In addition, the CPRR-Net method in pixel-level detection module was implemented under the deep learning open source framework TensorFlow. The tested frames were processed in computer with 16GB RAM, Intel (R) Xeon (R) Bronze CPU clocked at 2GHz, and GTX 2080Ti graphical processing unit (GPU) with 11 GB memory, 32-GB RAM.

B. TRACKING AND DETECTION EVALUATION

1) TRACKING PERFORMANCE

We compared several popular tracking networks, including Yolo, KCF and SiamFC based on the platform PYSOT. By using the same training dataset, we got the comparison results of these method in terms of accuracy and processing time for the testing dataset. We used the square bounding boxes in different size to focalize the contact region. To measure the performance of different methods, the mean center location precision (mCLP) was proposed. The mCLP is defined that the ratio of the number of contact point located in the inner box to the number of all the tracking images. The side length of the inner box was half of the proposed square bounding box, which purpose was to maintain the intersection feature of pantograph and catenary in the tracking region by an inner center-rectangle penalty mechanism.

In the Fig. 8-(a), the position of contact points was located in the inner yellow box, in which case the intersection feature of strip and catenary was strengthened in the green tracking box. By contrast the position of contact point was located outside of the yellow box in the Fig. 8-(b), so according to the definition of the mCLP, this was an invalid tracking.

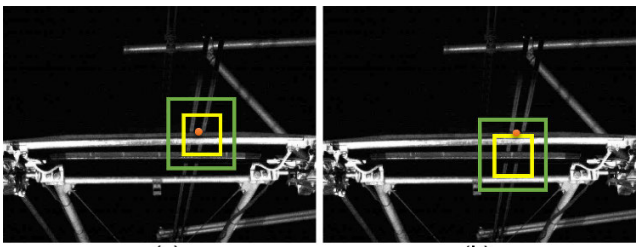


FIGURE 8. Definition of valid and invalid bounding boxes; the green box is the tracking box and the yellow box is the check box. (a) valid tracking; (b) Invalid tracking.

According to the new-defined evaluation criterion mCLP and mean processing time (mpT), the results of different method performance based on our new defined criteria in PAC scenario are presented in TABLE 3.

As shown in the TABLE 3, we found that Yolo performed not as well as SiamFC and KCF in terms of mean processing time. Furthermore, KCF had a better performance than SiamFC in mean center location precision, which is an essential indicator that we paid more attention in the scenario of PAC system. In the test, Yolo and SiamFC were tended to lose the target in our task while passing the location points

TABLE 3. The results of different tracking methods.

Method	Tracking Size	mCLP (%)	mpT (ms)
Yolo `	70×70	53.44	19.48
	100×100	78.73	20.72.
	150×150	84.86	22..38
KCF	70×70	64.78	10.69
	100×100	**97.43**	**11.52**
	150×150	91.2	12.87
SiamFC	70×70	79.05	9.92
	100×100	83.25	11.76
	150×150	92.4	14.1

or appearing of support beams during the tracking procession, whereas KCF possessed a robust tracking performance. Besides, the mpT of KCF meet the requirement of sampling frequency of 30fps. So, in this case, KCF was applied to locate the contact region in tracking section.

The collaboration of template matching and KCF tracking guarantees the consecutive tracking of contact regions. Fig. 9 illustrates the effectiveness of tracking method in distinguishing between messenger lines and contact lines and also in line-changing stage. Because messenger lines are farther than contact lines from camera and supplementary LED lights, contact lines are thicker and brighter in the acquired images. Fig. 9 (a) and (c) demonstrate that the matching and tracking method realized accurate location with the interference of messenger lines.

When passing the overlap span, as shown in Fig. 9 (b), the previous online contact line lifted up and the offline contact line went down to contact with the pantograph head. With suppletory light, the shadow of the pantograph head was dropped on the contact line because the contact line was not touched on the surface of the pantograph head. Therefore, based on our tracking method, template matching method was put into use to detect the online contact line in line-changing stage and the actually touching line was able to be distinguished.

In other complicated background, the tracking method was able to keep the accurate location of contact regions based on the line-intersection features and positions in the previous frame. For example, in Fig. 9 (d), the proposed tracking method kept tracking the contact region when the train was passing a cantilever and bridge structure.

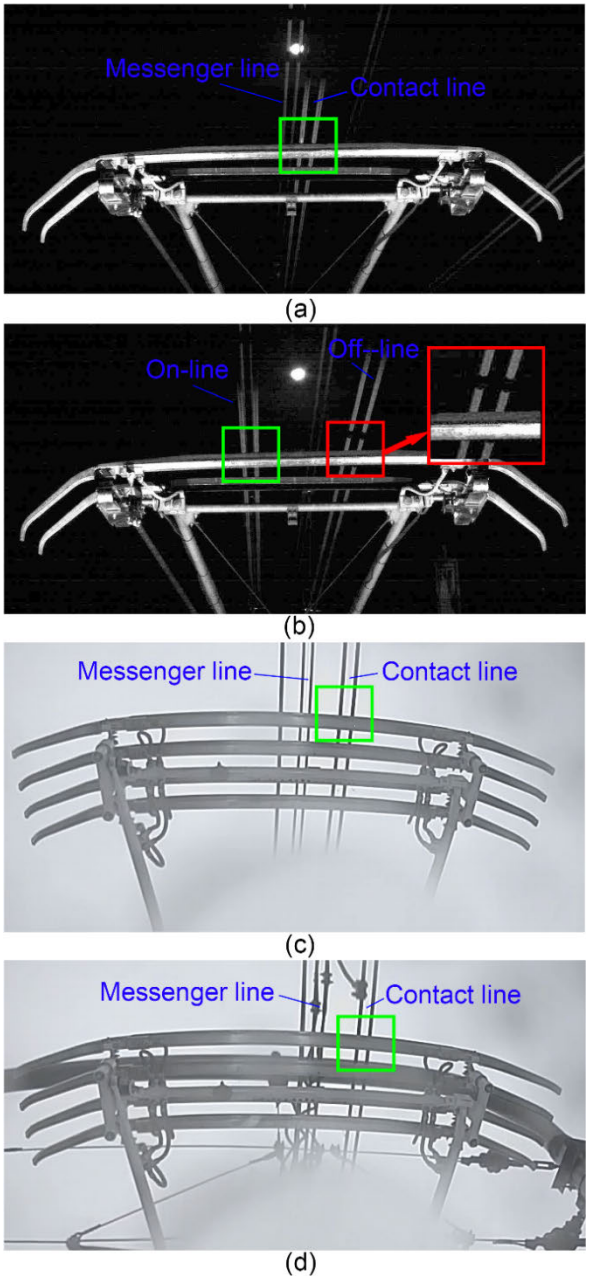


FIGURE 9. Tracking results in different scenarios. (a) Contact line and messenger in nighttime; (b) Changing of lines when passing though overlap span; (c) Contact line and messenger in daytime; (d) With complicated background.

2) CPT DETECTION PERFORMANCE

In this part, the added convolutional layer and the CPRR-Net were compared with other structures by a series of comparative tests.

In order to measure the accuracy rate of the CPT detection method based on our database, Coarse Positioning Region (CPR) and Fine Positioning Region (FPR) were defined [see in Fig. 10]. FPR and CPR were circles with a radius of 3 pixels and 5 pixels separately with the label position as the center. The detected position which located in FPR was defined as a true result according to the definition of FPR accuracy and, similarly the point in CPR was counted in CPR accuracy. Based on the database in this article, FPR accuracy requires precise detection results and CPR accuracy tolerates random distribution within a small range.

At the beginning of the network, a convolutional layer was added to smooth the input images and noise. TABLE 4 is the consequence of different type of filters and kernel sizes. For the FPR accuracy rate, both the two kernel sizes 3×3 and 5×5 have better performance compared with 7×7 size. The kernel size 5×5 performed better than 3×3 size in terms

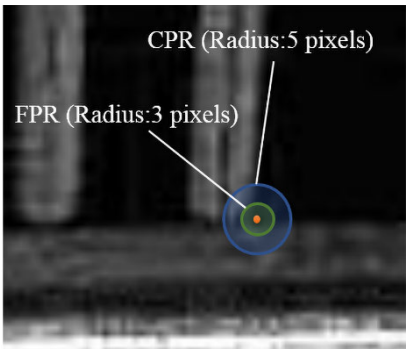


FIGURE 10. Definition of different regions for coarse positioning accuracy and fine positioning accuracy.

TABLE 4. The results of different filters.

Kernel size	FPR acc.	CPR acc
without	70.4	85.08
3×3	82.45	89.36
5×5	**85.61**	**94.50**
7×7	72.83	94.28

of CPR accuracy rate. Eventually the kernel size 5×5 was selected in the first convolutional layer.

The experimental results over all PAC images were summarized in TABLE 5. For comparison, we also reported results for several other structures, including AlexNet, VGG and GoogLeNet. Epoch set 5000 and all the method were convergent to a stable low number. Compared to different types of structure with performance of speed and accuracy, CPRR-Net can reach competitive performance with high accuracy. In order to optimize the structure of the CPRR-Net, different number of layers 18, 28 and 52 were compared. The performance of both CPRR-Net-52 and CPRR-Net-28 in the aspect of accuracy is similar and better than CPRR-Net-18. While for different number of layers, the more layers, the longer time needed for training and detection. Eventually CPRR-Net-28 was the best structure for this detection.

TABLE 5. The results of different filters.

Model	FPR acc (%)	CPR acc (%)	mpT (ms)
AlexNet	44.67	69.24	7.18
VGG	53.25	73.8	8.43
GoogLeNet	66.8	78.89	10.6
CPRR-Net-18	76.85	88.67	12.43
CPRR-Net-52	88.34	96.56	57.74
CPRR-Net-28A	80.62	93.4	15.22
CPRR-Net-28B	**85.61**	**94.50**	**15.32**
CPRR-Net-28C	86.87	94.34	17.53

In order to locate the position of CPT, a whole connected layer was attached and the comparison of different size of the last part was also listed in TABLE 5. CPRR-Net-28 A, B, C are separately1, 3, 5 layers in the whole connection part. The performance of type B and C was close, showing that the CPR accuracy was more that 94% and the FPR accuracy was more than 85%. However, more layers caused lower detection speed. Consequently 3 whole connection layers were selected considering both accuracy and real-time performance.

3) CPT CORRECTION PERFORMANCE

The Kalman filter was applied to correct and predict the CPT position by combining theoretical results and detected results. Fig.11 shows the CPT trajectory with and without optimization in the horizontal direction, and respectively demonstrates the errors from the ground truth. The detected positions are CPT coordinate detected by CPRR-Net and not yet optimized by the proposed filtering method. The optimized positions are the final results corrected with the combination of CPT motion equation and Kalman filter. Detection errors shown diminished within 3 pixels with optimization work. The comparison of errors showed the significant amelioration effect on the detection results, especially on large deviation area and singularities caused by arcing and other interference.

In zone 1, the error values of the detection results by CPRR-Net were almost distributed within the coarse positioning region. After the correction by the combination of contact point motion equation and Kalman filter, the errors were reduced to within 3 pixels and contracted to the fine positioning region, namely the accuracy of FPR was improved. In zone 2, the original detection results performed a large deviation from the calibration value, and the errors were within 5-10 pixels. After optimization, the error reduced within 5 pixels, which improved the accuracy of CPR. In zone 3, a singularity point appeared because the image was overexposed due to the severe arcing. The output of the CPRR-Net was far from the ground truth and the proposed optimization method reduced the error of the singularity within the CPR. Therefore, the proposed filter-based optimization method reduced greatly the errors and had a significant effect on the detection result. The statistics of the FPR accuracy and CPR accuracy in test database were given in TABLE 6.

TABLE 6. Accuracy (%) on detection results.

	CPRR-Net result	Optimization result
FPR	85.61	**97.07**
CPR	94.50	**99.97**

From the statistical results, the original results of CPRR-Net demonstrated the high performance in CPR accuracy, but the method also missed certain CPTs in more precise detection within 3 pixels. By the comparison before and after the optimization, the filter-based method in the last correction module improved more than 5 percent in CPR accuracy and more than 11 percent in FPR accuracy of the detected results. Conclusively, the combination of the CPT motion equation and Kalman improved both the accuracy and precision of the CPT position detection results.

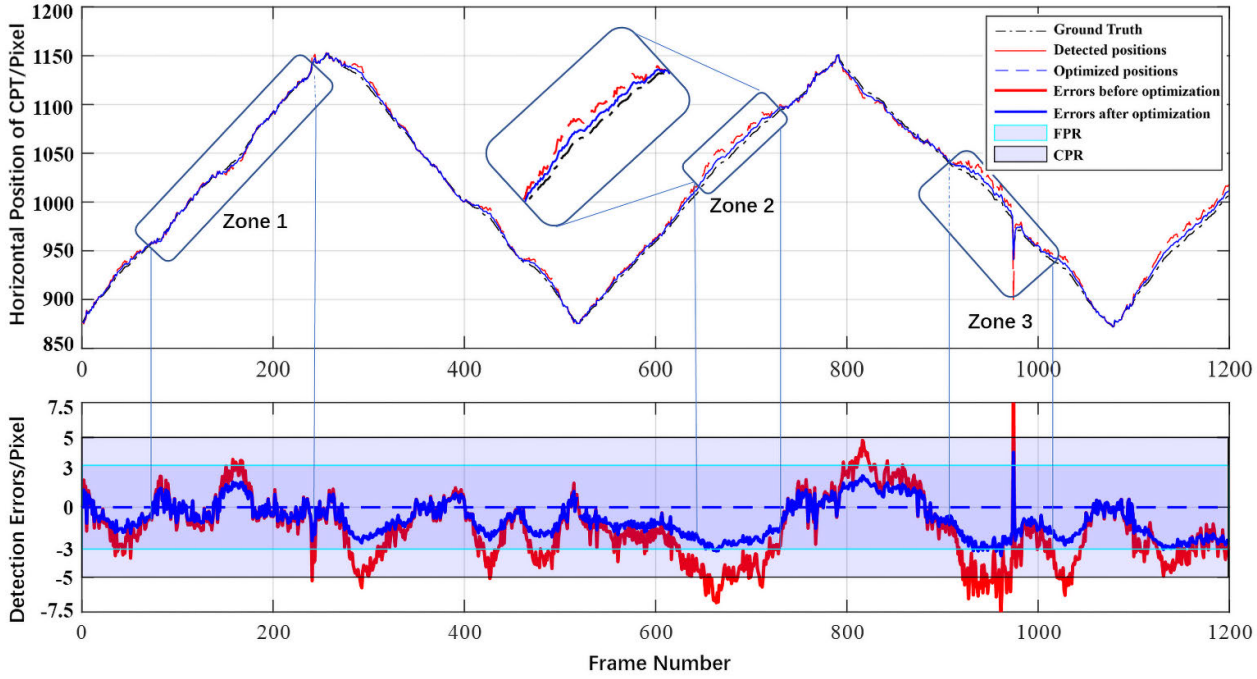


FIGURE 11. Demonstration of filter-based optimization results (CPT trajectory with and without optimization in the horizontal direction, and respectively demonstrates the errors from the ground truth).

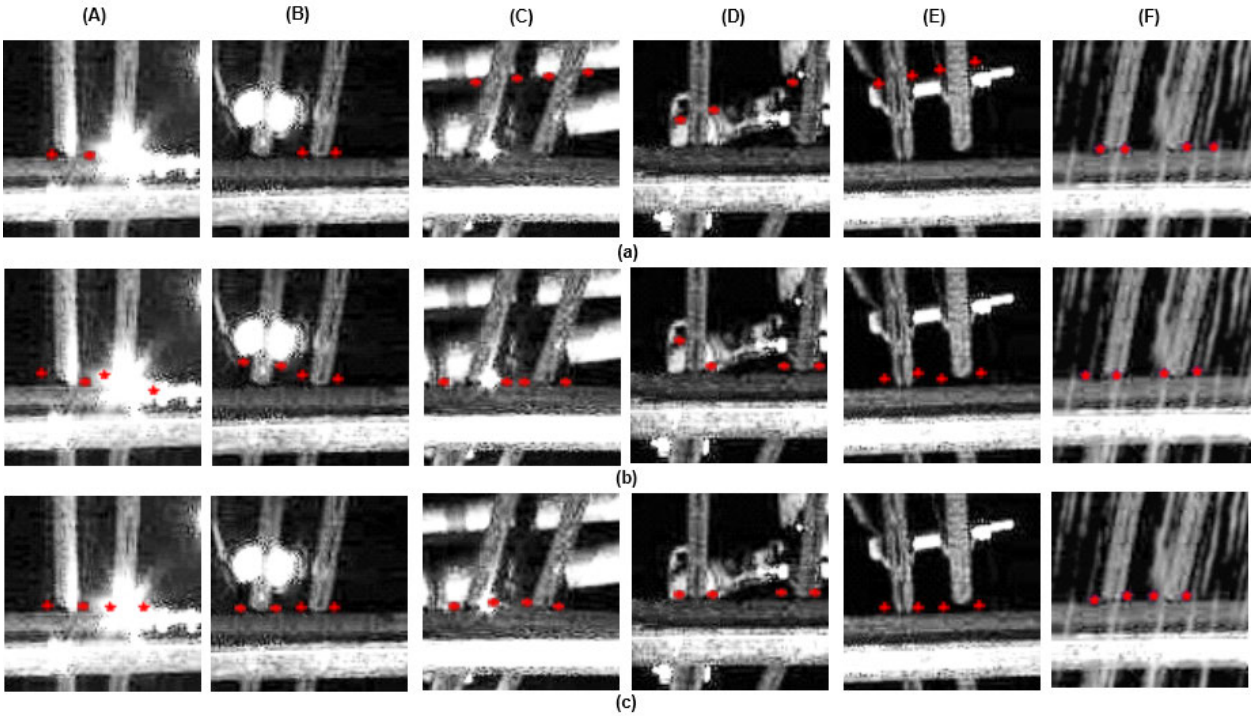


FIGURE 12. Comparison of detection results by different method: (a) Traditional method based on line detection; (b) The proposed method without Kalman filter; (c) The proposed method with Kalman filter, and by different interference: (A) Arcing, (B) Moon light, (C) Prop arm, (D) Location point, (E) Coupling bolt, (F) Unfavorable weather conditions.

C. THE CPT DETECTION RESULTS

For further validation of the proposed method, continuous railway data with various interference were adopted to verify the effectiveness and practicality. From the demonstration of detection results under different interference shown in images [see in Fig. 12], the comparison among the traditional method, the proposed method with and without Kalman filter corroborates the great improvement in CPT detection

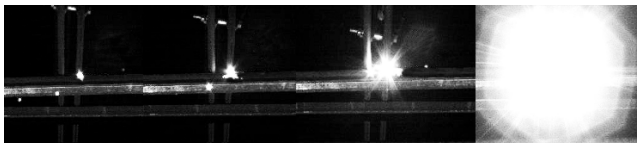


FIGURE 13. Consecutive record of arc evolution (interval was 20ms).

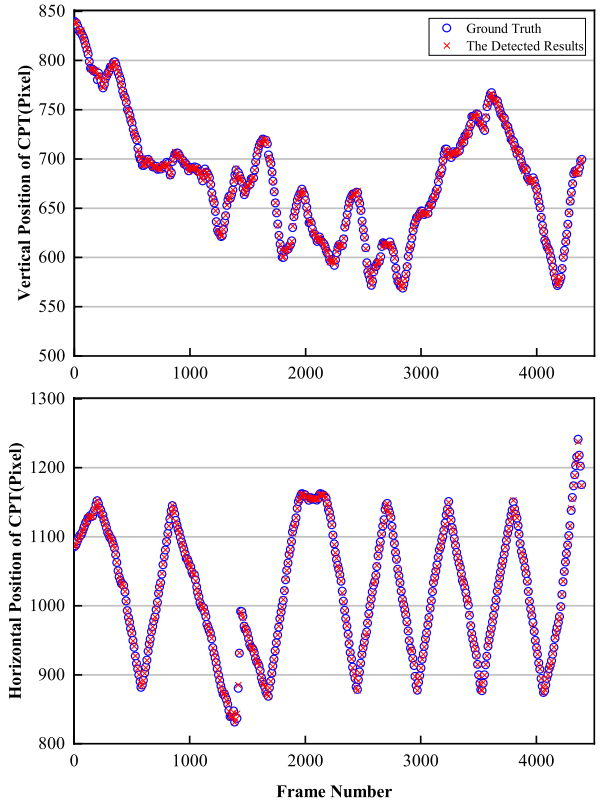


FIGURE 14. The detection result curves of the proposed method and ground truth in horizontal position and vertical position respectively.

performance with our new proposed method. In practical engineering application, detection results could be influenced by the interference of arcing, moon light, cantilevers, location points and coupling bolts and the contact points would be missed or false detected. For example, when arcing occurred, foreign light or unfavorable weather interfered and destroyed the straight-line feature of the contact line, the detection of certain CPT was missed [Fig. 12-(a)-A, B, F] by applying traditional line detection method. According to the criterion that the highest line in the image is the upper face of the pantograph slide, the contact points were incorrectly detected due to the presence of cantilevers [Fig. 12-(a)-C], location points [Fig. 12-(a)-D], coupling bolts [Fig. 12-(a)-E].

Based on the new proposed method in this article, the problems mentioned above can be solved methodologically and CPTs are all detected through overcoming the interference of complicated backgrounds. From the comparison of the proposed method without and with filter-based optimization method [Fig. 12-(b), (c)], the improvement of the detection performance is remarkable in view of that singularities caused by foreign lights and foreign objects were eliminated by combining the predicted coordinate and detected coordinate.

For consecutive arcing images [see in Fig. 13], the tracking module kept the correct tracking region even if one of frames was overexposed. Image information was lost and the results from CPRR-Net was deviated from the ground truth when severe arcing occurred, but with correction of the last optimization module, the deviation was eliminated within 5 pixels. It was also worth noting that the proposed method performed not well when there were long-time and severe arcing.

PAC-TPL2020 also adopted consecutive railway line images to verify the effectiveness of the proposed method. The overall CPT detection results in horizontal and vertical position were presented respectively in Fig.14 and it showed that the proposed method had the ability for accurate and precise detection of railway contact point. The experimental results proved that our proposed method was capable of consecutive detection and conductive to real-world railway monitoring with complicated interference.

V. CONCLUSION

Confronted with complicated backgrounds in real-world scenarios, a novel contact point detection method based on multiple strategies was proposed. The contact region was first tracked by an adapted KCF in real-time tracking module and then transmitted to the proposed CPRR-Net in a pixel-level detection module to get the CPT coordinate. The framework guarantees the robustness and instantaneity by reducing redundant background and data. Eventually, the detection results were optimized using Kalman filters and CPT motion equation in the last filter-based optimization module. This article also established a dataset PAC-TPL2020 which had more than 20,000 images with complicated background in real railway and ground truth of all the CPT coordinate. Experiments proved that the proposed method achieves a good performance in terms of accuracy, speed and robustness regardless of cantilever and bridge frame interference, unfavorable weather, arcing, foreign light interference, etc. In the following work, more challenges such as tunnel, long-lasting arcing interference, more complicated line conditions will be adopted to improve the richness of the PAC-TPL2020 and the robustness of the proposed method. Furthermore, based on the precise and high-speed CPT detection results, future work will focus on the calculation of contact force between the pantograph and the catenary.

REFERENCES

[1] W. Zhang, D. Zou, M. Tan, N. Zhou, R. Li, and G. Mei, "Review of pantograph and catenary interaction," *Frontiers Mech. Eng.*, vol. 13, no. 2, pp. 311–322, Jun. 2018, doi: 10.1007/s11465-018-0494-x.

[2] T. Koyama, M. Ikeda, S. Kobayashi, K. Nakamura, S. Tabayashi, and M. Niwakawa, "Measurement of the contact force of the pantograph by image processing technology," *Quart. Rep.*, vol. 55, no. 2, pp. 73–78, 2014, doi: 10.2219/rtriqr.55.73.

[3] E. Karakose, M. T. Gencoglu, M. Karakose, 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, doi: 10.1109/TII.2016.2628042.

[4] I. Aydin, M. Karakose, and E. Akin, "A new contactless fault diagnosis approach for pantograph-catenary system using pattern recognition and image processing methods," *Adv. Electr. Comput. Eng.*, vol. 14, no. 3, pp. 79–88, 2014, doi: 10.4316/AECE.2014.03010.

[5] A. Landi, L. Menconi, and L. Sani, "Hough transform and thermovision for monitoring pantograph-catenary system," *Proc. Inst. Mech. Eng., F, J. Rail Rapid Transit*, vol. 220, no. 4, pp. 435–447, Jul. 2006, doi: 10.1243/0954409JRRT41.

[6] 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, doi: 10.1109/SMC.2013.757.

[7] I. Aydin, M. Karakose, and E. Akin, "Monitoring of pantograph-catenary interaction by using particle swarm based contact wire tracking," in *Proc. 21st Int. Conf. Syst., Signals Image Process. (IWSSIP)*. Dubrovnik, Croatia: IEEE Computer Society, May 2014, pp. 23–26.

[8] I. Aydin, "A new approach based on firefly algorithm for vision-based railway overhead inspection system," *Measurement*, vol. 74, pp. 43–55, Oct. 2015, doi: 10.1016/j.measurement.2015.07.022.

[9] W. Zhou, H. Xiao, Z. Wang, L. Chen, and S. Fu, "Dynamic target template matching for railway catenary suspension motion detection in wind area," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 9, Sep. 2018, Art. no. 155014771879795, doi: 10.1177/1550147718797956.

[10] J. Lu, B. Liang, Q. Lei, X. Li, J. Liu, J. Liu, J. Xu, and W. Wang, "SCueU-net: Efficient damage detection method for railway rail," *IEEE Access*, vol. 8, pp. 125109–125120, 2020, doi: 10.1109/ACCESS.2020.3007603.

[11] B. Zhao, M. Dai, P. Li, R. Xue, and X. Ma, "Defect detection method for electric multiple units key components based on deep learning," *IEEE Access*, vol. 8, pp. 136808–136818, 2020, doi: 10.1109/ACCESS.2020.3009654.

[12] Y. Zhang, M. Liu, Y. Chen, H. Zhang, and Y. Guo, "Real-time vision-based system of fault detection for freight trains," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 5274–5284, Jul. 2020, doi: 10.1109/TIM.2019.2955799.

[13] D. H C, "An overview of you only look once: Unified, real-time object detection," *Int. J. for Res. Appl. Sci. Eng. Technol.*, vol. 8, no. 6, pp. 607–609, Jun. 2020, doi: 10.22214/ijraset.2020.6098.

[14] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, May 2017, pp. 7263–7271, doi: 10.1109/CVPR.2017.690.

[15] J. Redmon and A. Farhadi, "Yolov3," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1–6.

[16] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[17] L. Bertinetto, J. Valmadre, J. F. Henriques, A. Vedaldi, and P. H. S. Torr, "Fully-convolutional siamese networks for object tracking," in *Proc. Eur. Conf. Comput. Vis.*, Oct. 2016, pp. 850–865, doi: 10.1007/978-3-319-48881-3_56.

[18] Q. Guo, L. Liu, W. Xu, Y. Gong, X. Zhang, and W. Jing, "An improved faster R-CNN for high-speed railway dropper detection," *IEEE Access*, vol. 8, pp. 105622–105633, 2020, doi: 10.1109/ACCESS.2020.3000506.

[19] G. Karaduman and E. Akin, "A deep learning based method for detecting of wear on the current collector Strips' surfaces of the pantograph in railways," *IEEE Access*, vol. 8, pp. 183799–183812, 2020, doi: 10.1109/access.2020.3029555.

[20] X. Wei, S. Jiang, Y. Li, C. Li, L. Jia, and Y. Li, "Defect detection of pantograph slide based on deep learning and image processing technology," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 947–958, Mar. 2020, doi: 10.1109/TITS.2019.2900385.

[21] 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, doi: 10.1016/j.ins.2019.06.006.

[22] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, May 2015.

[24] C. Szegedy, "Going deeper with convolutions," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[26] D. Zhang, S. Gao, L. Yu, G. Kang, D. Zhan, and X. Wei, "A robust Pantograph–Catenary interaction condition monitoring method based on deep convolutional network," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 5, pp. 1920–1929, May 2020, doi: 10.1109/TIM.2019.2920721.

[27] Y. Shen, Z. Liu, and G. Zhang, "PAC interaction inspection using real-time contact point tracking," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 10, pp. 4051–4064, Oct. 2019, doi: 10.1109/TIM.2018.2884039.

[28] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 583–596, Mar. 2015, doi: 10.1109/TPAMI.2014.2345390.

[29] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, May 2015.

XUAN YANG received the B.S. degree in engineering mechanics from Southwest Jiaotong University, in 2018. He is currently pursuing the engineer's degree with the Ecole Centrale de Nantes and the master's degree with SWJTU.

His research interests include computer vision, deep learning, mechanical modeling, and simulation.

NING ZHOU received the B.S. and M.S. degrees from Xi'an Jiaotong University and the Ph.D. degree from Southwest Jiaotong University. He is currently an associate researcher with the State Key Laboratory of Traction Power, Southwest Jiaotong University, China.

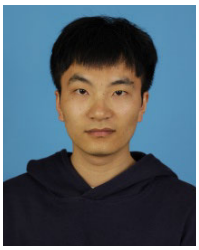
His current research interests include nonlinear dynamics, active control, signal processing, and computer-aided engineering.

YUEPING LIU received the M.S. degree in computer application technology from Northeast Petroleum University in 2013. She is currently pursuing the Ph.D. degree with the State Key Laboratory of Traction Power, Southwest Jiaotong University.

Her research interests include machine learning and pattern recognition.

WEI QUAN (Member, IEEE) received the B.S. and M.S. degrees in control theory and engineering and the Ph.D. degree in signal and information processing from Southwest Jiaotong University (SWJTU), in 2004, 2007, and 2013, respectively. He is currently an Associate Professor of control science and engineering with SWJTU.

His research interests include computer vision, AI and deep learning, and image processing.



XUEMIN LU (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering and its automation from East China Jiaotong University, Nanchang, China, in 2016. He is currently pursuing the Ph.D. degree in pattern recognition and intelligent system with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu, China. In 2020, he was selected into the Sakura Science Plan funded by the Japan Science and Technology Agency, and received the State Scholarship Fund from the China Scholarship Council and will become a joint Ph.D. Student with the Department of Electrical and Computer Engineering, University of Alberta, Canada. His current research interests include AI, deep learning, computer vision, and automatic detection systems.

WEIHUA ZHANG received the B.Sc. degree in mechanical automation, the M.Sc. degree in engineering mechanics, and the Ph.D. degree in vehicle operation engineering from Southwest Jiaotong University, Chengdu, China, in 1983, 1989, and 1996, respectively. He is currently a Professor with the State Key Laboratory of Traction Power, Southwest Jiaotong University. His research interests include rail vehicle design theory and structural reliability, vehicle system dynamics and control, pantograph-catenary relation, and measurement and test technology.