A Novel Low-Light Catenary Image Enhancement Approach for CSCs Detection in High-Speed Railways

WEIPING GUO, HUI WANG (Student Member, IEEE), ZHIWEI HAN (Member, IEEE), JUNPING ZHONG (Member, IEEE), AND ZHIGANG LIU (Senior Member, IEEE)

School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China

Corresponding Author: Z. Liu (e-mail: liuzg_cd@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51977182.

ABSTRACT The catenary system is essential for ensuring the stable energy transmission of trains in high-speed railways. The non-contact catenary detection is a promising monitoring method, where the catenary image is captured by an industrial camera mounted on the inspection vehicle. The image quality is susceptible to the limitations of the environment and equipment, which adversely affects the catenary location and detection accuracy. In this paper, we propose an unsupervised learning-based catenary image enhancement method for improving localization accuracy. First, the enhancement model is optimized to enhance the catenary image quality, making it sharper and more conducive to detection. Subsequently, an advanced small target location approach, called TPH-YOLOv5, is used to locate the catenary components. Finally, we compared the localization performance of the enhanced image with the low-light image. The experiment results show that the proposed method can effectively enhance the quality of low-light catenary images and improve the positioning accuracy.

INDEX TERMS Catenary, high-speed railway, image enhancement, component localization, unsupervised learning.

I. INTRODUCTION

CATENARY is an essential part of the high-speed railway, which provides electricity to the trains. As shown in Fig. 1, the catenary support devices (CSCs) consist of many components, such as the insulator, brace sleeve, double sleeve connector, etc. And each component includes many smaller parts, such as nut, base, and screw pin. These components do not function individually but are combined to form a complex system [1]. If an abnormality or failure occurs in one component, it is likely to cause effectiveness or even collapse of the entire system. To make matters worse, the working environment of the catenary is poor, which is also one of the reasons for the abnormal occurrence of the catenary. Therefore, it is essential to study a method to improve the efficiency of catenary inspection.

Computer image processing is a commonly used catenary detection method, currently. As shown in Fig. 2, catenary images are taken on the railway by the railway inspection vehicle running. The quality of the image and the shooting environment have a significant impact on the effectiveness of this method. Due to the trains frequently running during the day and the background of the catenary under sunlight being complex, the detection vehicle generally runs at night. However, poor ambient light conditions at night make it easy to cause insufficient brightness of the captured image. These dark data present degraded image features, vague outlines, low contrast, and poor visibility [2]. These factors lead to ineffective extracting of feature information during catenary detection, reducing detection efficiency and accuracy. Therefore, it is essential to propose a method to improve the image quality for better image detection.

At present, some image enhancement methods are applied in catenary detection. Generally, these methods can be divided into two categories: traditional image processing methods and deep learning-based methods. For example, Wu et al. [2] proposed an approach based on wavelet-based
GUO et al.: NOVEL LOW-LIGHT CATENARY IMAGE ENHANCEMENT APPROACH FOR CSCs DETECTION

FIGURE 1. Catenary images taken by the catenary inspection system. Some components are indicated: a) Insulator base, b) Insulator, c) Brace sleeve, d) Rotary double ear, e) Steady arm base, f) Bracing wire hook, g) Double sleeve connector, h) Messenger wire base, i) Isoelectric line. Marked in blue are some small components: j) Screw pin, k) Screw pin fix, l) Screw U-nut, m) Screw nut.

FIGURE 2. Catenary image acquisition system. The catenary inspection vehicle collects image data as it runs on the railway. As shown in the red dotted box at the bottom left of the figure, the camera and light source system are installed on the top of the railway inspection vehicle to fill in light and take catenary images at night.

contourlet transform, which can eliminate the redundancy of contourlet transform and the pseudo-Gibbs effects. Wu et al. [3] used nonsubsampled contourlet transform with an adaptive enhancement function to enhance the catenary image. Chen et al. [4] utilized the retinex theory for removing the noise to achieve catenary image enhancement. However, although these methods have specific effects, they generally have the problems of poor generalization and slow processing speed. In recent years, methods based on deep learning have made significant progress in catenary image enhancement. Chen et al. [5] proposed a multi-feature fusion network to learn coarse and fine details by combining latent image features. Liu et al. [6] proposed a network that connects the convolutional neural network (CNN) and 3-D lookup tables theory to enhance the catenary images. These deep learning-based methods are more demanding on the dataset, and it is difficult to obtain the lack of light and dark paired catenary images to compose paired datasets that can be used for convolutional neural network training in practice. Therefore, some images need to be generated artificially to form the data of light and dark pairs. However, there is a difference between the manually generated data and the real-world images, which will lead to poor results for the model.

In response to the above problems, we propose a catenary image enhancement model, which does not require paired datasets, avoiding the problems caused by artificial data, and whose processing is better and faster. The overall workflow of this method is exhibited in Fig. 3. The main contributions of this study are summarized as follows:

1) First, a model of catenary image enhancement based on the Zero-Reference Learning for Low-Light Image Enhancement (Zero-DCE) [7] network is proposed, which is unsupervised learning and does not require paired datasets. Compared with other state of art image enhancement networks, the processing speed of this method is faster and the processing result is better.

2) Second, the accuracy of positioning is improved after enhancing the catenary data. The target detection experiments with the enhanced data from the first step and the original low-light images demonstrate that the localization accuracy is significantly improved after enhancement.

3) Third, the improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios (TPH-YPLOv5) [8] is adopted to carry out the location of small catenary components. This method is more effective in the field of small object positioning. Through comparative experiments, the prospect of this method in the positioning of small parts of the catenary is shown.

This paper is organized as follows. Section II shows the methodology of the proposed enhancement method. Section III presents the experimental results and discussion. Finally, conclusions and further improvements of this work are drawn in Section IV.

II. METHODOLOGY

To carry out catenary detection better, we proposed an improved enhancement method based on the Zero-DCE network [7] that makes the model smaller, inference faster, and performing effectively. As shown in Fig. 3, this network consists of three parts: a sparse block (SB), a feature attention block (FAB), and a curve enhancement block (CEB). Significantly, the SB module and the FAB are specifically designed to estimate a set of optimal Light-Enhancement curve (LEC) parameter maps used in the CEB. And then, the CEB iteratively processes the low-light image through the LEC to achieve enhancement.

A. DEEP CURVE ESTIMATION NETWORK

To obtain better LECs, we modify the DCE network structure, adding the SB and FAB module to the network. Inspired by the theories [9], [10], we added dilated and standard convolution layers in SB to expand the receptive field size, which improves network performance without increasing model parameters greatly. And then, in FAB, both local and global
features are integrated with a skip-connection to enhance the expressiveness of the LEC. A set of eight LEC parameter maps generated by FAB is used by CEB to enhance the low-light image. The LEC is shown in Eq. (1), which iteratively enhances the low-light images.

$$L_{En}(x) = LE_{n-1}(x) + A_n(x)LE_{n-1}(x)(1 - LE_{n-1}(x))$$  \hspace{1cm} (1)

where \( n \) denotes the number of iterations, which is set to eight after experiments. \( A \) is the LEC parameter map with the same size as the input image. And the \( x \) refers pixel of the input image.

### B. LOSS FUNCTIONS

To realize unsupervised learning, Zero-DCE [7] network designs four sets of unique loss functions with no reference. We adopt three of these loss functions, which are described as follows, taking into account the color insensitivity of catenary detection.

First, the Exposure Control Loss \( (L_{exp}) \) is designed to control exposure levels and calculates the difference between the average intensity value of the local area and the exposure. The parameter \( E \) is set to work well. It works best in catenary image data when we set \( E = 0.29 \). The loss \( L_{exp} \) is defined as:

$$L_{exp} = \frac{1}{M} \sum_{k=1}^{M} |Y_k - E|$$  \hspace{1cm} (2)

where \( M \) refers to the number of local regions with the size of \( 16 \times 16 \). \( Y \) means the value of mean intensity in the local areas.

Second, the Spatial Consistency Loss \( (L_{spa}) \) is designed to preserve spatial consistency by reducing the difference in the neighborhood regions between the input image and the enhanced image. The loss \( L_{spa} \) is defined as follows:

$$L_{spa} = \frac{1}{K} \sum_{i=1}^{K} \sum_{j \in \Omega(i)} \left( |Y_i - Y_j| - |I_i - I_j| \right)^2$$  \hspace{1cm} (3)

where \( K \) represents the number of local regions with the size of \( 4 \times 4 \). \( \Omega(i) \) refers to the four adjacent regions of the upper, lower, left, and right, which is centered at region \( i \). Finally, \( Y \) is used to represent the average intensity of the local regions in the enhanced image. \( I \), similar to \( y \), represent the input image.

Third, the Illumination Smoothness Loss \( (L_{tvA}) \) establishes monotonicity constraints between adjacent pixels. The \( L_{tvA} \) are as follows:

$$L_{tvA} = \frac{1}{N} \sum_{n=1}^{N} (|\nabla_x A_n| + |\nabla_y A_n|)^2$$  \hspace{1cm} (4)

where \( N \) represents the number of iterations, which is set to eight here. \( \nabla_x \) and \( \nabla_y \) represent the gradients in the horizontal and vertical directions, respectively.

Finally, the total loss is defined as:

$$L_{total} = W_{tvA}L_{tvA} + W_{spa}L_{spa} + W_{exp}L_{exp}$$  \hspace{1cm} (5)

where \( W_{tvA} \), \( W_{spa} \), \( W_{exp} \) denote the loss weighting factor. We determine the value of the loss weighting factor by a set of parameter search experiments, setting the \( W_{tvA} = 1500 \), \( W_{spa} = 1.5 \), and \( W_{exp} = 5 \), respectively.

### III. ANALYSIS AND DISCUSSION OF EXPERIMENTAL RESULTS

To evaluate the performance of the enhancement method proposed in this paper, a series of model training and testing experiments are carried out based on the PyTorch deep learning framework [11]. Next, we compare some other enhancement methods to prove the effectiveness of our model. Finally, we introduce the TPH-YOLOv5 [8] object detection method into the field of catenary detection. The enhancement effect is verified by comparing the positioning accuracy of the dataset enhanced or not, at the same time, the performance of this catenary small target localization method is measured.

**Dataset:** As shown in Fig. 2, the experimental data is composed of catenary images, all of which are taken at night by the CCD high-speed camera installed on the roof of the railway inspection vehicle. A total of 4000 catenary image were obtained, of which 3301 and 300 suitable images were used to train and test the enhancement model, respectively. In addition, 2000 images were selected for testing the TPH-YOLOv5 object detection model, and 634 images were used to test. The resolution of the images collected by the railway inspection vehicle is \( 6600 \times 4400 \). To facilitate model training, the size of these images is scaled \( 600 \times 400 \).

**Training Process:** The experiment environment is described as follows: NVIDIA TITAN RTX GPU with 24 GB memory, Intel Xeon E5-2640 V4 CPU with 32 GB RAM, and
Ubuntu18.04 system. The parameters of the networks are initialized with standard zero mean and 0.02 standard deviation Gaussian function. The ADAM optimizer is chosen as the backpropagation gradient descent method with a fixed learning rate 0.0001 and weight decay 0.001. The batch size and the max epoch are set to 16 and 100, respectively.

A. EVALUATION METRICS
Some metrics include the mean squared error (MSE), peak signal-to-noise ratio (PSNR), structural similarity (SSIM), brightness, and contrast are introduced as image quality evaluation indexes. Among these indicators, the higher the value of PSNR, the better the quality of the processed image and the less noise, and the higher the value of SSIM, the more similar the processed image is to the image before processing. In addition, MSE can evaluate the degree of change in the data. The smaller the value of MSE, the better the accuracy of the experimental data obtained by the model. Contrast is critical to visual effects, for example, high contrast is very beneficial for image clarity and detail performance. Finally, the Frames Per Second (FPS) is used to indicate the computational speed, which represents the number of frames that can be processed per second.

\[
\text{MSE} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} (X(i,j) - Y(i,j))^2 \quad (6)
\]

\[
\text{PSNR} = 10 \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (7)
\]

\[
\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{\mu_x^2 + \mu_y^2 + C_1(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (8)
\]

\[
\text{Brightness} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} x^2(i,j) \quad (9)
\]

\[
\text{Contrast} = \sum_{\delta} \delta(i,j)^2 P_\delta(i,j) \quad (10)
\]

where \(H\) and \(W\) are the height and width of the image respectively, \(X(i,j)\) is the enhanced image, \(Y(i,j)\) is the reference image. \(\text{MAX}_I\) is the maximum value of the image value range, which is 255 if each sample point is represented by 8 bits. \(\mu_x\) and \(\mu_y\) are the average of the image, \(\sigma_x^2\) and \(\sigma_y^2\) are the image variances, and \(\sigma_{xy}\) is the image covariance. \(\delta(i,j)\) renders the grayscale difference between adjacent pixels, and \(P_\delta(i,j)\) represents the probability of pixel distribution with a grayscale difference \(\delta\) between adjacent pixels.

B. ABLATION STUDY
To demonstrate the effectiveness of each part of this network, the ablation studies were carried out as follows. Different network structures of our method are tried and we present the results of different module combinations in Fig. 4. It can be seen that the enhancement effect without the SB module is not obvious, and it is not much different from the input image. When the FEB module is missing, the enhancement result is noticeable, but the detail and contrast are not as good as the full result. This means that the modules are effective in enhancing the image and improving the details.

In addition, quantitative analysis using different structures is presented in Table 1. It is clear that the results are the same as those in Fig. 4. Our method had the best performance, reaching 33.32 PSNR and 32.81 MSE, followed by the lack of FEB, and the lack of SB module the worst. Finally, we expanded the convolutional kernel size of each convolutional layer from 32 to 64 and found that the results improved by 0.74 PSNR and 2.76 MSE, but its model weight expanded from 349.5 KB to 1.3MB, and the inference speed became much slower, from 25.98 FPS to 19.71 FPS. In practical application, the network is generally used in mobile inspection vehicles. Due to the limitations
of mobile miniaturized equipment, the compute capability is insufficient. To achieve fast and efficient processing, we chose 32 convolutional kernels of each layer.

C. QUANTITATIVE COMPARISON OF DIFFERENT METHODS

To compare the performance of our model, we used two representative methods called Dual Illumination Estimation for Robust Exposure Correction (DUAL) [12] and Lighting Network for Low-Light Image Enhancement (DLN) [13]. As shown in Table 2, we can see that our method has the highest values of PSNR, SSIM, and moderate values of brightness and contrast. The effects of different enhancement models are compared in Fig. 5. It can be seen from the figure that the DUAL [12] has too much noise and minor enhancement. In addition, the enhancement results of DLN [13] have too low contrast and appear as a phantom shadow. On the contrary, our method performs better for natural exposure and precise details.

D. LOCALIZATION OF CATENARY IMAGES

To compare the performance of image enhancement for detecting small components of the catenary, the following experiments based on the TPH-YOLOv5 [8] were conducted. A total of 4,000 catenary pictures were obtained for this experiment. Screening of these images resulted in 634 low-light images or about 16% of the total. We count the average grayscale values of these 634 low-light images and determine the image enhancement threshold. Finally, we selected 2000 suitable images as the training set, 200 as the validation set, and 634 low-light images as the test set.

Three loss functions, including box-loss, objectness-loss, and classification-loss, are used by the TPH-YOLOv5 [8] model, which respectively represents the distance between the prediction frame and the ground truth (GIoU), the confidence of the network, and the classification correctness of the anchor frame and the ground truth. According to the data set and the hyperparameters selected by the model, the epoch time is set to 300. As shown in Fig. 6, we can see that the loss function gradually converges after 200 epochs. Meanwhile, in Fig. 7, the values of the precision (P), recall (R) and mean average precision (mAP) gradually become stable after 150 epochs. Therefore, the epoch selected in this experiment can meet the training requirements. The mathematical expression of R, P, AP and.

| Algorithm | PSNR↑ | SSIM↑ | Brightness | Contrast |
|-----------|-------|-------|------------|----------|
| Our method | 33.32 | 0.96  | 40.72      | 149.90   |
| DLN       | 25.95 | 0.78  | 57.20      | 111.19   |
| DUAL      | 15.93 | 0.58  | 35.21      | 180.32   |

FIGURE 5. Result comparisons on different typical methods. Red boxes point out the detailed differences between the catenary components.

FIGURE 6. Trends of the loss functions for 300 epochs.

FIGURE 7. Trends of the positioning effect evaluation indexes for 300 epochs.
TABLE 3. Average precision for small components localization.

| Categories                              | Low-light images | enhanced images |
|-----------------------------------------|------------------|-----------------|
| Pieces of porcelain rod insulator       | 0.985            | 0.995           |
| U-nut of porcelain rod insulator        | 0.976            | 0.983           |
| Bracing wire                            | 0.821            | 0.861           |
| Bracing wire hook                       | 0.734            | 0.743           |
| Binaural sleeve                         | 0.971            | 0.979           |
| Screw nut                               | 0.941            | 0.948           |
| Wire of screw nut                       | 0.821            | 0.861           |
| Hook of messenger wire base             | 0.859            | 0.886           |
| Slot of messenger wire base             | 0.919            | 0.943           |
| Average Precision                       | 0.930            | 0.941           |

mAP are shown below.

\[
R = \frac{TP}{TP + FN} \quad (11)
\]

\[
P = \frac{TP}{TP + FP} \quad (12)
\]

\[
AP = \int_{0}^{1} PRdr \quad (13)
\]

\[
mAP = \frac{1}{N} \sum_{i=1}^{N} AP_i \quad (14)
\]

where TP (True positive) represents the number of actual positive samples that are predicted as positive, FP (False positive) means the actual number of negative samples predicted to be positive, and FN (False negative) indicates the number of actual positive samples that are predicted to be negative. Meanwhile, N denotes the number of categories in the dataset, which is set to 40 here.

After obtaining a valid localization model by the above experiments, we perform object detection inference on 634 low-light images and calculate their localization accuracy. Then, we enhance the test set using the enhancement model mentioned above. Finally, the target detection inference is performed on the enhanced dataset obtained in the previous step. The experimental results are shown in Fig. 8 and Table 3. In Fig. 8, due to the enhancement process, the small components of the catenary are correctly located, whereas some parts in the low-light image are missed or errors detected. For example, in Fig. 8 (b), the red arrows indicate the missed detection components in the low-light image. The majority of the improved component positioning accuracy is improved in Table 2. In a few cases, the localization accuracy is degraded, which may be due to the loss of enhanced features caused by partially blurred imaging.

Finally, to compare the performance of TPH-YOLOv5, we selected the state-of-the-art object detection models, Swin-Transformer [17] and YOLOv5 [20], to conduct comparative experiments. The training parameters and the appropriate number of iterations recommended by the original author were adopted, and the same dataset and experimental environment were used. Table 4 shows the localization results for the catenary small components dataset with different methods. It can be seen that the Precision (P) and mAP of our method achieved the best grades of 0.941 and 0.604. However, after using Swin-Transformer, the results become 0.851 and 0.487 which is the worst performance of the three methods. At the same time, the result of the YOLOv5 method is in the middle, reaching 0.922 and 0.583. The scale range of catenary components may be too large, which leads to the poor practicability of the model for this application.

IV. CONCLUSION

With the improvement of detection technology, the quality of catenary images has gradually become one of the bottlenecks limiting the performance of CSCs detection. In summary, the main contributions of this paper are as follows. We propose a low-light image enhancement method that significantly
optimizes the enhancement effect. The small component localization method was adopted to compare the localization accuracy of low-light and enhanced images. Experimental results verified the enhancement of the catenary image is beneficial to the subsequent localization. Finally, the localization results also show the superiority of this object detection method for positioning small components of the catenary. In the future, we will continue to improve the network structure and will verify the effectiveness of the enhancement for anomaly detection for CSCs.

REFERENCES

[1] Z. Liu, Y. Song, Y. Han, H. Wang, Z. Han, and J. Zhang, “Advances of research on high-speed railway catenary,” J. Southwest Jiaotong Univ., vol. 51, no. 3, pp. 495–518, Jun. 2016.

[2] C. Wu, Z. Liu, and H. Jiang, “Catenary image enhancement using wavelet-based contourlet transform with cycle translation,” Optik, vol. 125, no. 15, pp. 3922–3925, 2014.

[3] C. Wu, Z. Liu, and H. Jiang, “Choosing the filter for catenary image enhancement method based on the non-subsampled contourlet transform,” Res. Sci. Instrum., vol. 88, no. 5, pp. 453–476, 2017.

[4] Y. Chen, B. Song, Y. Zeng, X. Du, and M. Guizani, “A deep learning-based approach for fault diagnosis of current-carrying ring in catenary system,” Neural Comput. Appl., to be published.

[5] Y. Chen, B. Song, X. Du, and N. Guizani, “The enhancement of catenary image with low visibility based on multi-feature fusion network in railway industry,” Comput. Commun., vol. 152, pp. 200–205, Feb. 2020.

[6] W. Liu et al., “An automatic loose defect detection method for catenary bracing wire components using deep convolutional neural networks and image processing,” IEEE Trans. Instrum. Meas., vol. 70, pp. 1–14, 2021, doi: 10.1109/TIM.2021.3113121.

[7] C. Guo et al., “Zero-reference deep curve estimation for low-light image enhancement,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2020, pp. 1780–1789.

[8] X. Zhu, S. Lyu, X. Wang, and Q. Zhao, “TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios,” in Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops, 2021, pp. 2778–2788.

[9] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei, and H. Liu, “Attention-guided CNN for image denoising,” Neural Netw., vol. 124, pp. 117–129, Apr. 2020.

[10] 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.

[11] J. Zhong, Z. Han, Y. Han, and W. Zhang, “A CNN-based defect inspection method for catenary split pins in high-speed railway,” IEEE Trans. Instrum. Meas., vol. 68, no. 8, pp. 2849–2860, Aug. 2019.

[12] Z. Liu, W. Liu, and Z. Han, “A high-precision detection approach for catenary geometry parameters of electrical railway,” IEEE Trans. Instrum. Meas., vol. 66, no. 7, pp. 1798–1808, Jul. 2017.

[13] Z. Liu, Y. Nie, and W. S. Zheng, “Dual illumination estimation for

[14] L.-W. Wang, Z.-S. Liu, W.-C. Siu, and D. P. K. Lun, “Lightening et al.

[15] G. Zhang, Z. Liu, W. Liu, and Z. Han, “Non-contact detection of conductors height & stagger of contact line based on camera calibration,” J. China Railway Soc., vol. 36, no. 2, pp. 10012–10022.

[16] A. Paszke et al., “PyTorch: An imperative style, high-performance deep learning library,” in Advances in Neural Information Processing Systems, vol. 32. Red Hook, NY, USA: Curran, 2019.

[17] Q. Zhang, Y. Nie, and W. S. Zheng, “Dual illumination estimation for robust exposure correction,” Comput. Graph. Forum, vol. 38, no. 7, pp. 243–252, 2019.

[18] X. Zhu, S. Lyu, X. Wang, and Q. Zhao, “TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios,” in Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops, 2021, pp. 2778–2788.

[19] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei, and H. Liu, “Attention-guided CNN for image denoising,” Neural Netw., vol. 124, pp. 117–129, Apr. 2020.

[20] 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.

[21] Y. Lyu, Z. Han, J. Zhong, C. Li, and Z. Liu, “A generic anomaly detection of catenary support components based on generative adversarial networks,” IEEE Trans. Instrum. Meas., vol. 69, no. 5, pp. 2439–2448, May 2020.

WEI-PING GUO received the B.S. degree in electrical engineering and its automation from the Henan University of Science and Technology, Luoyang, China, in 2019. He is currently pursuing the M.Sc. degree with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu, China.

His current research interests include image processing, machine learning, computer vision, and their application in railway fault detection.

HUI WANG (Student Member, IEEE) received the B.S. degree in electrical engineering and its automation from the University of South China, Hengyang, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu, China.

His current research interests include image processing, deep learning, computer vision, and active pantograph control.

ZHWEI HAN (Member, IEEE) received the Ph.D. degree in power system and its automation from the Southwest Jiaotong University of China in 2013.

He is currently a Lecturer with the School of Electrical Engineering, Southwest Jiaotong University. His current research interests include modern signal processing, computer vision, and their application in railway and electric power system.

JUNPING ZHONG (Member, IEEE) received the B.S. degree in electronic and information engineering from Southwest Jiaotong University, Chengdu, China, in 2014, where he is currently pursuing the Ph.D. degree in electrical engineering.

His current research interests include image processing, computer vision, and their applications in railway fault detection.

ZHIGANG LIU (Senior Member, IEEE) received the Ph.D. degree in power systems and its automation from Southwest Jiaotong University, China, in 2003. He is currently a Full Professor with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu. He is also a Guest Professor with Tongji University, Shanghai. He has authored three books and published more than 200 peer-reviewed journal and conference articles. His research interests include the electrical relation of EMUs and traction, detection, and assessment of pantograph-catenary in high-speed railways.

Prof. Liu is an Associate Editor-in-Chief of IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT and an Associate Editor of IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and IEEE ACCESS. He received the IEEE TIM's Outstanding Associate Editors for 2019, 2020, and 2021, and the Outstanding Reviewer of IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT in 2018.