Adversarial Reconstruction Based on Tighter Oriented Localization for Catenary Insulator Defect Detection in High-Speed Railways

Wang, H.; Nunez, Alfredo

DOI
10.1109/TITS.2020.3020287

Publication date
2020

Document Version
Final published version

Published in
IEEE Transactions on Intelligent Transportation Systems

Adversarial Reconstruction Based on Tighter Oriented Localization for Catenary Insulator Defect Detection in High-Speed Railways

Junping Zhong, *Member, IEEE*, Zhigang Liu, *Senior Member, IEEE*, Cheng Yang, *Student Member, IEEE*, Hongrui Wang, *Member, IEEE*, Shibin Gao, and Alfredo Núñez, *Senior Member, IEEE*

***Abstract*— The catenary insulator maintains electrical insulation between catenary and ground. Its defects may happen due to the long-term impact from vehicle and environment. At present, the research of defect detection for catenary insulator faces several challenges. 1) Localization accuracy is low, which causes the localized object to be incomplete or/and merge with unnecessary background. 2) Horizontal localization brings inevitable unnecessary information because horizontal box cannot fit well with the shape of insulator. 3) Supervised learning models for defects recognition are unreliable as the available defect samples are insufficient to train models well. To address these issues, this article proposes a novel two-stage defect detection method. In the localization stage, a novel localization network called TOL-Framework is constructed to reduce the background and realize tighter oriented localization. Compared with general basic framework Faster R-CNN, the TOL-Framework cascades a regression module inside basic framework and adds an external postprocess network, which is adversarially trained by standard insulators to refine the localization. These two novel steps greatly improve the oriented localization accuracy. In the defect detection stage, an adversarial reconstruction model that is trained only using normal samples is proposed to evaluate the defect states. A comparison with other methods is conducted using a dataset collected from a 60km section of the Changsha-Zhuzhou railway line in China. The results show the proposed method has the highest localization accuracy, and is effective for insulator defect detection.**

***Index Terms*— Catenary insulator, tighter oriented localization, generative adversarial network, defect detection, deep learning, high-speed railways.**

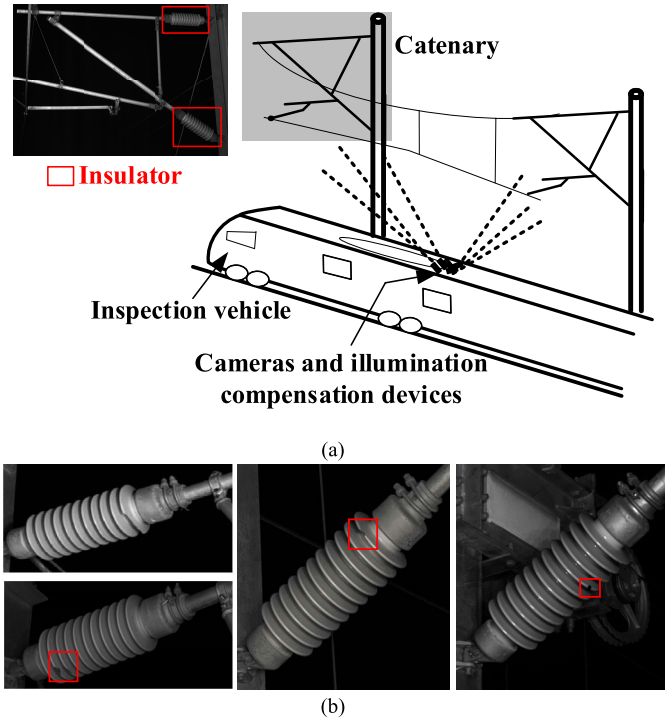


Fig. 1. Catenary inspection vehicle and defects of insulators. (a) Sketch map of the catenary of the inspection vehicle and the position of the insulator. (b) Defective insulators.

I. Introduction

CATENARY plays an important role in transmitting electric power from the contact line to the train, and the reliability can directly affect railway transportation. As a key part of the catenary, the insulator is not only a supporting and connecting component but also maintains electrical insulation between the catenary and the ground. Due to the long-term mechanical impact triggered by vehicles and the complex environment along the rail line, insulators can become damaged, which poses a risk to railway operation safety. Therefore, it is essential to inspect the defective states of insulators.

Over the past few years, railway infra-managers have tried to replace inefficient on-site manual inspection with computer vision-based methods for inspecting catenary components [1]. As shown in Fig. 1(a), multiple high-resolution cameras and illumination compensation devices are mounted on the roof of an inspection vehicle. Images are captured at night to avoid complex illumination and background. The cameras are mounted symmetrically to ensure that both sides of insulators can be inspected. The red bounding boxes in Fig. 1(a) show the

Manuscript received October 14, 2019; revised March 26, 2020 and July 5, 2020; accepted August 13, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 51977182 and Grant U1734202. The Associate Editor for this article was Z. Duric. *(Corresponding author: Zhigang Liu).*

Junping Zhong, Zhigang Liu, Cheng Yang, and Shibin Gao are with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China (e-mail: zhongjunping@my.swjtu.edu.cn; liuzg_cd@126.com; yangc@my.swjtu.edu.cn; gao_shi_bin@126.com).

Hongrui Wang and Alfredo Núñez are with the Section of Railway Engineering, Delft University of Technology, 2628 Delft, The Netherlands (e-mail: soul_wang0@163.com; a.a.nunezvicencio@tudelft.nl).

Digital Object Identifier 10.1109/TITS.2020.3020287

insulator positions in the catenary, and some defective samples are shown in Fig. 1(b). To accomplish defect detection, two cascade stages are always needed. The object localization stage is first applied to the global image, and the position of the insulator is predicted with a bounding box. Then, in the defect detection stage, the defective states are recognized on the localized insulator. However, inaccurate localization makes the localized insulator incomplete or enclosed with background, which directly interferes with the defect detection stage. In addition, defective samples cannot generally be collected in practice, and the quantity of available defective samples makes it difficult to construct defect detection models. Therefore, effective image processing algorithms for component localization and defect detection in catenary inspection systems are urgently needed.

For object localization, traditional methods based on handcrafted features, such as SIFT (scale-invariant feature transform) [2], HOG (histogram of oriented gradient) [3] and LBP (local binary patterns) [4], have been widely used in railways [5]–[7] and other fields [8], [9]. Zhong *et al*. [5] applied template matching on a standard catenary sleeve image and an original image to search the object position based on SIFT. Cho and Ko [6] applied a variant of SIFT that is robust to changes in camera viewpoint to locate pantograph positions. Han *et al*. [7] used cascade support vector machines to classify a series of sliding window images, which are represented by a HOG to realize the localization of the catenary clevis. Fan *et al.* [8] proposed a line LBP encoding method to represent a target object, which could be used to localize fasteners on the rail track when the illumination is complex. In [9], [10], HOG-based deformable part models (DPMs) were adopted to localize pedestrian and human faces. However, most traditional methods are not robust to changes in size, shape, illumination, position, background of the object, etc. In the case of catenary, when traditional methods craft certain features to detect the insulators, they can perform well on one type of images in which attributes are relatively fixed. When those attributes change, the performance of traditional methods declines [5], [23], [24]. With deep learning, we have more parameters to be tuned than with traditional methods, giving more flexibility to capture various real-life catenary and measurement conditions.

In recent years, CNN-based deep learning techniques have shown great power in object localization, and they follow two kinds of basic frameworks: R-CNN (region-based convolutional neural network) [11] and YOLO (You Only Look Once) [12]. Most recently proposed localizers are based on R-CNNs. Ren *et al.* [13] proposed Faster R-CNN, which can accelerate the proposal generation and realize end-to-end training. He *et al.* [14] proposed a residual network (ResNet) to train deeper backbone networks and extract more discriminative features than other networks. Dai *et al.* [15] proposed the R-FCN (region-based fully convolution networks), which adopts a fully convolutional network as a feature extractor to accelerate the deep ResNet. Lin *et al*. [16] constructed an FPN (feature pyramid network) by exploiting the inherent multiscale and pyramidal hierarchy of deep convolutional networks. In [17], a multistage extension of R-CNN called cascade R-CNN was proposed to improve the localization quality. Some localizers are also proposed based on YOLO. Liu *et al*. [18] presented a single-shot multibox detector (SSD) that regresses a set of default bounding boxes on multiple layers to realize real-time localization. Redmon and Farhadi [19] introduced the faster and stronger architecture YOLO9000, which can achieve a trade-off between speed and accuracy. However, all these methods generally localize the object with a horizontal bounding box, and only a few works [20]–[22] have investigated arbitrarily oriented localization based on an R-CNN. Both types of deep learning methods based on horizontal bounding boxes have been proposed for the localization of catenary components [23]–[26]. However, these methods face difficulties when fitting the shape and the position of insulators, making the localized insulator incomplete or enclosed with unnecessary background.

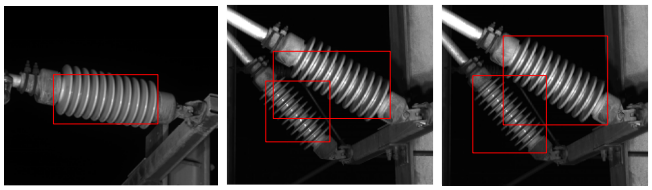


Fig. 2. Insulators localized by current methods. (a) Incomplete localization. (b) Incomplete localization and unnecessary background information. (c) Unnecessary background information.

For defect detection, Karakose *et al*. [27] used a Hough transform and edge extraction to detect the surface defects of a pantograph effectively. Supervised classification methods were applied in [23], [28], [30], [31] to classify different states of railway fasteners or pantographs directly. However, supervised learning is not suitable for the detection of defective insulators because usually large number of defective samples are unavailable. This makes it difficult, for instance, to effectively train neural networks. Few unsupervised learning methods, such as autoencoders [32] and GAN (generative adversarial network) [33], have been used for surface defect detection of catenary insulators [25], [34] and other abnormality detection tasks [35], [36]. The existing catenary insulator detection methods [25], [34] are based on horizontal localization, which is not effective for defect detection.

Overall, various problems from currently available methods need to be solved for defect detection of catenary insulators. Some of those problems as listed below.

1) The localization accuracy is not high, which causes the insulators to not all be localized. Those localized insulators can be incomplete or include unnecessary background information, as shown in Figs. 2(a) and 2(b).

2) Horizontal bounding box localization yields inevitable unnecessary background information even if horizontal bounding boxes are perfectly predicted, as shown in Fig. 2(c).

3) There is a lack of defective samples. In practice, it is hard to collect large numbers of different defective samples to construct a defect detection model.

To address these problems, the proposed method in this article considers three neural networks, namely BLN, GAN-1 and GAN-2. First, we use BLN to localize insulators with

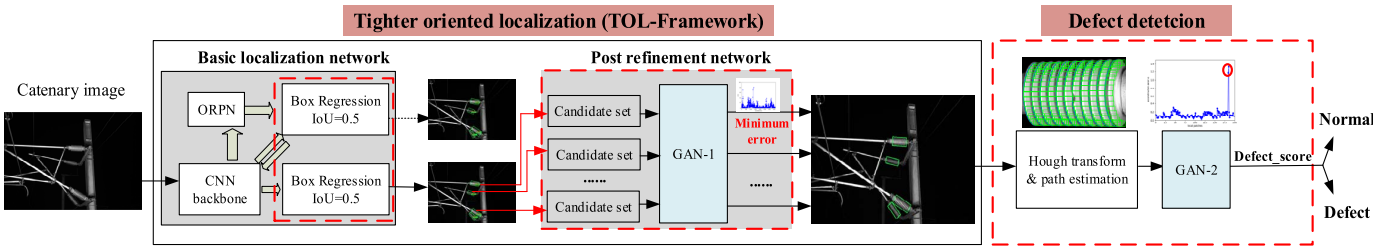


Fig. 3. Overview of the proposed insulator defect detection method.

oriented boxes rather than using traditional horizontal boxes. Then, the GAN-1 is applied to further adjust the localized box to enclose insulator tightly. Finally, the GAN-2 is utilized to detect defects on the localized insulators. The GAN-2 is introduced because it is trained only using normal samples, which solves the issue that the available defect samples are insufficient to construct defect detection models.

We summarize the contributions of this article as follows.

1) A novel TOL-Framework (BLN with GAN-1) is proposed to realize arbitrarily oriented localization and greatly improve the accuracy compared with several competitive methods. The novelties of the TOL-Framework are listed below.

- The angle information of the insulator is used for bounding box regression, which can realize oriented localization.
- A cascade regression module inside the generic Faster R-CNN framework is proposed to refine the oriented bounding box preliminarily.
- To further improve the localization accuracy, we included a GAN (called GAN-1) behind the basic localization network. The GAN-1 selects the bounding box image which has the largest similarity with the insulator as the final localized bounding box, which makes the obtained box enclose the insulator better. It is a new application of GAN for improving localization accuracy.

2) A new defect detection method based on a GAN (called GAN-2) is proposed for catenary insulators, whose advantages are as follows:

- The training of GAN-2 only needs normal samples. This solves the issue that the available defect samples are insufficient to construct defect detection models.
- The reconstruction error is computed in local insulator patches, which have a relatively fixed texture and size. Thus, even small defects can be detected effectively.

The rest of this article is organized as follows. Section II provides an overview of the insulator defect detection method. Section III introduces the structure of the proposed TOL-Framework and describes how to realize tighter oriented localization. Section IV presents the details of the defect detection process. Section V gives the experiments to evaluate the performance of our method. Finally, conclusions are drawn in Section VI.

II. Overview of Our Method

The overview of our method is shown in Fig. 3, which mainly consists of two successive stages, namely, tighter oriented localization and defect detection. The image size is 6, 600×4, 400 pixels, which ensures that defects in the insulator can be clearly observed even though the defective regions are very small.

A. Tighter Oriented Localization

To solve the inevitable problems caused by the horizontal-box localization presented in Fig. 2 and to further improve the localization accuracy, a TOL-Framework (tighter oriented localization framework) is proposed in this article. The workflow of the TOL-Framework is described as follows. Original catenary images are input into a basic localization network, which includes the *CNN backbone*, *ORPN* and a *cascaded oriented box regression*, to predict a preliminary oriented box for each target insulator. Then, for each preliminary bounding box, a candidate set can be produced according to the extreme IoU, which ensures that nearly all insulators are localized. The candidate sets are input into a GAN-1 module, which is trained by only using normal standard insulator images to compute the reconstruction error. Finally, the candidate that has the minimum error is regarded as the localized bounding box for this insulator.

Compared with the structure of the current basic localization model [13], [20], [21], the TOL-Framework cascades a bounding box regression module inside the model and adds an external postprocessing module to refine the positions of bounding boxes predicted by the basic model. In addition, a rotational RPN is applied, and a pyramid feature with ResNet-101 as the CNN backbone is selected to realize oriented localization. Details of the TOL-Framework are described in Section III.

B. Defect Detection

The accurate oriented localization stage has provided insulator images that are tightly enclosed nearly without unnecessary background information, and all insulators can be rotated to the horizontal and have similar textures, which is helpful for defect detection. As the available defective samples are insufficient for deep model training, we propose a GAN-based method that only uses normal samples for training. The workflow of defect detection is as follows. An HT&PE (Hough transform & path estimation) algorithm is applied on the localized insulator to extract insulator patches through the body of the insulator. Then, all the patches are input into GAN-2, which is trained by only using normal insulator patches to compute the reconstruction error. Finally, defects and their

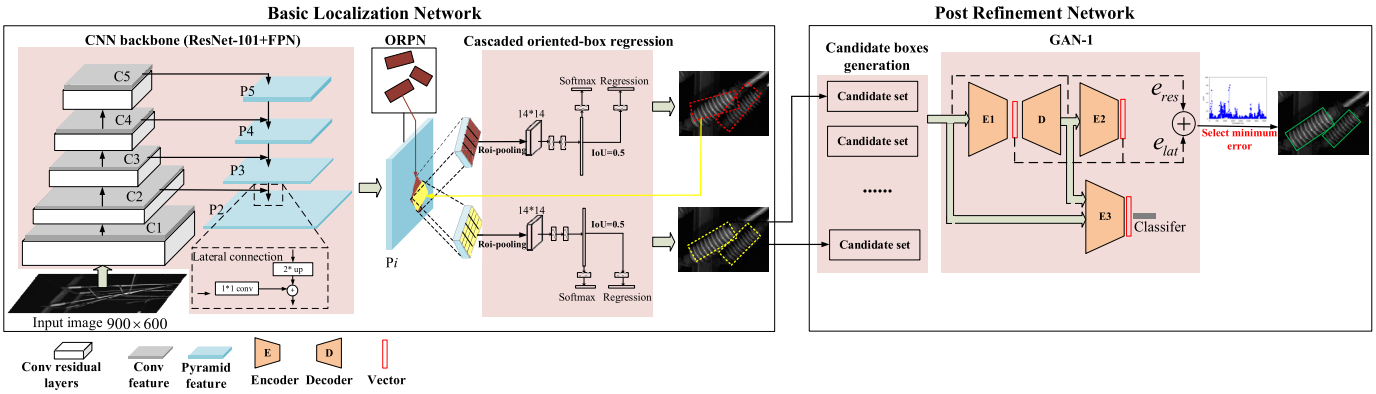


Fig. 4. Structure of the TOL-Framework.

positions are evaluated by a defect score (*Defect_score*), which is defined according to the reconstruction error. Details of defect detection are presented in Section IV.

III. Tighter Oriented Localization

In this section, we describe the structure of the proposed TOL-Framework and explain how it realizes tighter oriented localization. The TOL-Framework includes two parts, namely, a basic localization network (BLN) and a post-refinement network (PRN), as shown in Fig. 4.

In general, the existing catenary localization method follows the following architecture: *CNN backbone->RPN->horizontal-box regression*. In this article, a new architecture is proposed, namely, *CNN backbone->ORPN->cascaded oriented bounding box regression*, and it is used as a basic localization network in the TOL-Framework to predict preliminary bounding boxes of insulators. To further improve the oriented localization accuracy and obtain tighter boxes, a post-refinement network behind the basic localization network is added. The details of the TOL-Framework are elaborated as follows.

A. Basic Localization Network

The BLN follows the widely used R-CNN architecture [13], [20], [21], and it consists of three subnetworks, namely, the *CNN backbone*, *oriented region proposal network* (ORPN) and *cascaded oriented bounding box regression*. The CNN backbone extracts semantic features from the input catenary image first. Then, the ORPN generates oriented proposals for the insulator targets based on the extracted feature maps. Finally, two regression modules are successively applied to produce the preliminary oriented box for the insulator.

1) CNN Backbone (ResNet-101+FPN):: The CNN backbone processes the input image and generates conv feature maps at multiple levels. Unlike the commonly used methods [11]–[15], [17], [20]–[24] that only use the top-level feature map, we build a feature pyramid network (FPN) [16] and use pyramid features as the extracted features. The FPN is effective as each pyramid level is fused with higher-level semantic features of the backbone. For the backbone, a widely used ResNet model [14] with 101 conv layers is chosen. Specifically, as shown in Fig. 4, the captured 6, 600 × 4.400 pixels catenary image is resized to 900 × 600 pixels, which goes through bottom-up conv-residual blocks of ResNet and produces activation maps at five scales with a scaling step of 2 by feedforward computing. The generated features of the backbone are defined as {C1, C2, C3, C4, C5}, which are the last output features of each conv-residual block. These features (except for C1 due to its large memory consumption) are further enhanced from the top-down pathway via lateral connections. Each lateral connection upsamples the upper pyramid feature by a factor of 2 and merges with the corresponding bottom-up map, which adds a 1 × 1 convolutional layer to reduce the channel dimensions. The enhanced pyramid features {P2, P3, P4, P5, P6} are the final extracted features of the CNN backbone.

2) Oriented Region Proposal Network: An RPN [13] is a sliding window class agnostic object detector. A traditional RPN applies an anchor strategy, which has scale and aspect ratio parameters, to the top-level feature map to produce initial horizontal bounding boxes, which are called anchors. These anchors are classified and regressed as proposals. Compared with a traditional RPN, two adjustments are made to design the ORPN. First, an orientation parameter is added to the anchor strategy to produce rotated proposals for insulators. Second, proposals are produced not only on the top-level feature map but also on the pyramid features. Specifically, the workflow of the ORPN can be described as follows. At each point of the feature map, Pi (i = 2, 3, 4, 5, 6), 49 (7 × 7 × 1, orientations × aspect rations × scale) oriented anchors are produced with 7 orientations{0°, 15°, 30°, 45°, 60°, 75°, 90°}, 7 aspect ratios {1/5, 1/3.5, 1/2, 1, 2, 3.5, 5} and one scale $[32\times(i-1)]^2$. Then, the oriented anchors are sent to the object/background classification module and regression is coordinated to generate proposals. Finally, the proposals are mapped to the corresponding regions in the feature map Pi, as shown by the red bounding box in Pi in Fig. 4, and passed to the cascaded oriented bounding box regression subnetwork.

3) Cascaded Oriented Bounding Box Regression: Generally, only a single bounding box regression module trained with IoU = 0.5 is applied. As the top box regression shows in Fig. 4, the feature of the proposal is downsampled to 14×14 pixels by ROI pooling and is sent to several fully connected

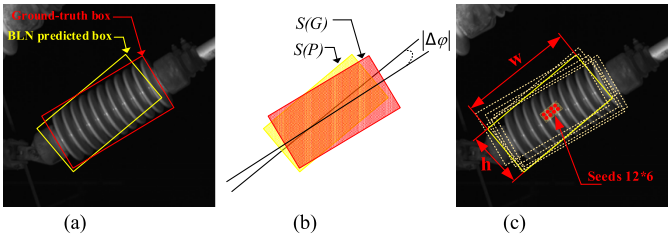


Fig. 5. Candidate bounding box generation. (a) Ground-truth bounding box and the bounding box predicted by the BLN. (b) Overlap and angle differences. (c) Seed region and candidate bounding boxes.

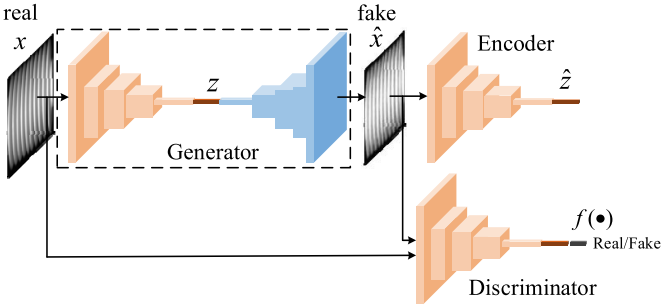


Fig. 6. Generative adversarial network.

layers for classification and regression. Inspired by a cascade R-CNN [17], a cascaded oriented bounding box regression module, which is trained with IoUs of 0.5 and 0.5, is applied to improve the localization accuracy. Three or more regression modules are not cascaded because there is little accuracy improvement when more regression modules are applied. The reason for setting the second IoU to 0.5 rather than a higher IoU is that many boxes output by the first regression will be filtered due to their low IoUs, and training samples will not be enough for the added regression module. As shown in Fig. 4, the yellow bounding boxes predicted by the cascaded regression modules are much closer to the true positions of the insulators when compared with the red bounding boxes predicted by single regression.

B. Post-Refinement Network

Even though the localization accuracy of the BLN is obviously improved, it still cannot satisfy the requirement of defect detection. In this article, we propose a post-refinement network based on a defined reconstruction error to make the localized boxes tighter to their true positions, which is computed by a trained GAN (generative adversarial network). We show the overall process in ***Algorithm 1***. The details of the two subprocesses are also presented below.

Algorithm 1 Post-Refinement

1: **Input:** m Boxes localized by the BLN $\{B_1, B_2, \ldots, B_m\}$
2: **Output:**m Post-refined boxes $\{B'_1, B'_2, \ldots, B'_m\}$
3: **for** each box $B_i (i <= m)$ **do**
4: Apply *candidate bounding box generation* and obtain a candidate bounding box set $\{B_{i1}, B_{i2}, \ldots, B_{in}\}$
5: **for** $j <= n$ **do**
6: send B_{ij} into the trained GAN-1 to compute reconstruction error e_{ij}
7: **end for**
8: select the candidate bounding box which has the minimum loss $min\{e_{i1}, e_{i2,}\ldots, e_{in}\}$ as the post-refined bounding box B'_i
9:**end for**

1) Candidate Bounding Box Generation: In Section V.B, we will prove that all the boxes produced by BLN have IoUs > 0.8 & $|\Delta\varphi| < 10°$(overlaps with their true positions are more than 0.8, and angle differences with their true angles are less than 10°) in the testing experiment. The definitions of IoU [21] and $|\Delta\varphi|$ are shown in formula (1) and Fig. 5(b). Therefore, for each bounding box B_i predicted by the BLN, a minimum region that contains the central points of potential ground-truth bounding boxes can be formed. In this article, we extend the minimum region and define a seed region that has the same central point as B_i and has a size of [w/5, h/5]. The seed region is divided into $12\times$ seeds, as shown in Fig. 5(c). In each seed, 50 candidate boxes (with 5 scales of 0.9, 0.95, 1, 1.05, 1.1 based on [w, h], and with 10 angles by an angle stride of 2° in [−10° ∼10°], thus $5 \times 10 = 50$) will be generated. Finally, the seed region can generate 3,600 ($12 \times 6 \times 50$) candidate boxes, which form the candidate set $\{B_{i1}, B_{i2}, \ldots, B_{in}\}$ in ***Algorithm 1***. Note that each candidate set includes the ground-truth bounding box or a bounding box that is very close to the ground-truth bounding box.

$$IoU = \frac{S(\mathrm{P}) \cap S(\mathrm{G})}{S(\mathrm{P}) \cup S(\mathrm{G})} \tag{1}$$

2) Generative Adversarial Network Refinement: Horizontal standard catenary insulators have similar appearances. In each candidate set $\{B_{i1}, B_{i2}, \ldots, B_{in}\}$, we can select the bounding box that has the largest similarity with the standard insulator as the final localized bounding box. Thus, we utilize a GAN-1 model to evaluate the similarity. The GAN-1 model was adversarially trained by only using a large amount of standard normal insulators, and it only fits well with a standard insulator. The overview of the GAN-1 model is shown in Fig. 6, which includes a generator (encoder-decoder) and a discriminator (encoder followed by a softmax classifier $f(\bullet)$). The structures of the encoder and decoder follow the DCGAN in [37]. During training, the contextual loss, adversarial loss and latent loss are computed, which are expressed in formulas (2)-(4), respectively. The main training objective is to minimize a weighted sum of these losses, as shown in formula (5). Minimizing the contextual loss indicates that the generator makes the texture of the generated image $\hat{x}$ similar to the texture of the input x. Moreover, minimizing the adversarial loss and latent loss makes the discriminator unable to classify x (the real image) and $\hat{x}$(the generated fake image) correctly.

Contextual loss:

$$L_{con} = E_{x\sim p_x} \left| x - \hat{x} \right|_1 \tag{2}$$

Adversarial loss:

$$L_{adv} = E_{x\sim p_x}[\log D(x)] + E_{x\sim p_x}[\log(1 - D(\hat{x}))] \tag{3}$$

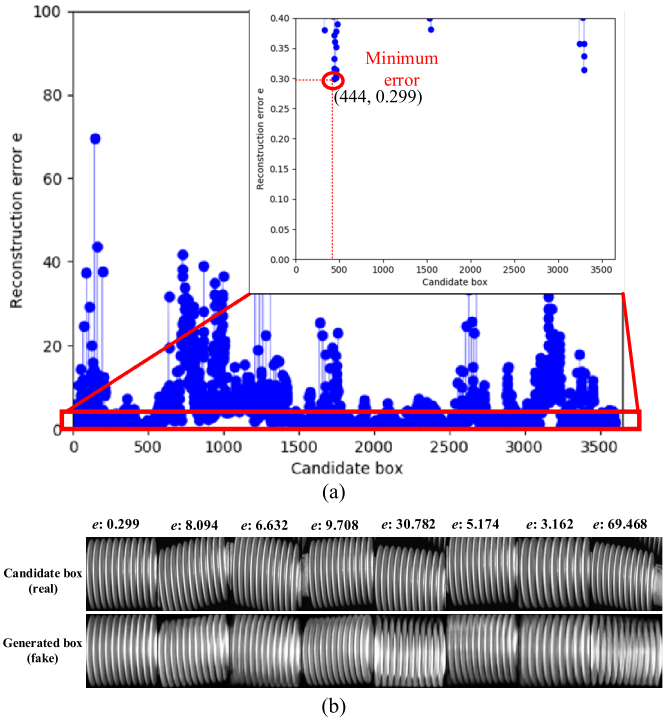


Fig. 7. GAN-based post-refinement. (a) The box with the minimum reconstruction error is selected. (b) Comparison between different candidates and their fake images.

Latent loss:

$$L_{lat} = E_{x\sim p_x} \left| f(x) - f(\hat{x}) \right|_2 \tag{4}$$

Total loss

$$L_{train} = L_{adv} + 30 * L_{con} + L_{lat} \tag{5}$$

Note that the GAN-1 model is trained by only using standard insulators, and the trained parameters of the GAN-1 model cannot fit well with nonstandard insulators. When a nonstandard insulator and a standard insulator are tested in the GAN-1 model, the standard insulator has a smaller contextual error and latent error than the nonstandard insulator. Thus, we define the reconstruction error e, defined in formula (6), to evaluate the difference between the test image and the standard insulator image. For a bounding box B_i predicted by the BLN, when its candidate bounding boxes $\{B_{i1}, B_{i2}, \ldots, B_{in}\}$ are input into the trained GAN-1 model and their reconstruction errors are computed, the candidate box with minimum error is regarded as the post-refined bounding box. Fig. 7(a) shows an example of the reconstruction errors computed on a candidate set. The 444th candidate bounding box with the minimum error of 0.299 was selected as the post-refined bounding box. Fig. 7(b) shows some candidate bounding boxes and their fake images produced by the generator. It shows that the fake image of the candidate bounding box, which is similar to the standard insulator, changes minimally when compared with the candidate, and the reconstruction error is also minimal.

$$e = 2 * L_{con} + L_{lat} \tag{6}$$

IV. Defect Detection Based on a Generative Adversarial Network

To address the issue of insufficient defective samples for defect detection, we utilize a GAN-2 model, which was trained on normal insulator patches to obtain the reconstruction errors of the test patches; defects can be detected according to the obtained errors. The structure of the GAN-2 model in this part is similar to that of the model GAN-1 used in localization in section. III.B, but we apply the GAN-2 model on insulator patches rather than global insulators because the difference in the reconstruction errors between normal global insulators and defective global insulators is not obvious, which makes small defects undetectable. In addition, the defect position cannot be confirmed when the GAN is applied on a global insulator. In this part, HT&PE (Hough transform and path estimation) is first used to extract insulator patches throughout the body of the insulator, and then the extracted patches are input into the GAN to distinguish states.

A. Local Insulator Patch Extraction by HT&PE

Suitable insulator patches should have similar textures to help the GAN-2 model detect defects and cover the whole body of the insulator to ensure that each part of the insulator is detected. To obtain this kind of insulator patch, we apply an HT (Hough transform) to obtain the top and bottom points of the insulator and estimate suitable sliding paths. Then, a window with a fixed scale slides along the paths and crops the local insulator images as extracted patches. The details of the HT&PE are described as follows.

As the localized insulator is rotated horizontally, the top and bottom points of the insulator pieces are nearly collinear, as shown by the red bounding boxes in Fig. 8(b). Therefore, these two groups of points can be obtained by using an HT (Hough transform) [38], which detects two straight lines in the angle ranges $(-90°, 80°]$ and $[80°,90°)$. Fig. 8(e) shows the Hough matrix produced by HT. The points in the red boxes are the two maximum values that correspond to two groups of points in Fig. 8(b). The definition of angle θ is shown in Fig. 8(a). Moreover, the midpoints (yellow boxes (x_1, y_1) and (x_2, y_2) in Fig. 8(b)) and the average distance r between adjacent insulator pieces can be obtained. Therefore, we can roughly evaluate the window-sliding paths with the following ellipse formula.

$$\begin{cases} x = x_{cent} + 0.5 * a * \cos(\mathrm{t}) \\ y = y_{cent} + r * \sin(\mathrm{t}), \end{cases} \quad \mathrm{t} \in (\frac{\pi}{k}, \frac{2\pi}{k}), k = (1, 2, \ldots, 30) \tag{7}$$

where (x_{cent}, y_{cent}) is the midpoint of (x_1, y_1) and (x_2, y_2) and a is the distance between (x_1, y_1) and (x_2, y_2), which is regarded as the long axis of the ellipse. Therefore, a window with a fixed width of $1.2*r$ slides along the evaluated paths through the body of the insulator to obtain insulator patches, as shown in Figs. 8(c) and (d). Note that the patches that exceed the region between two detected lines are removed.

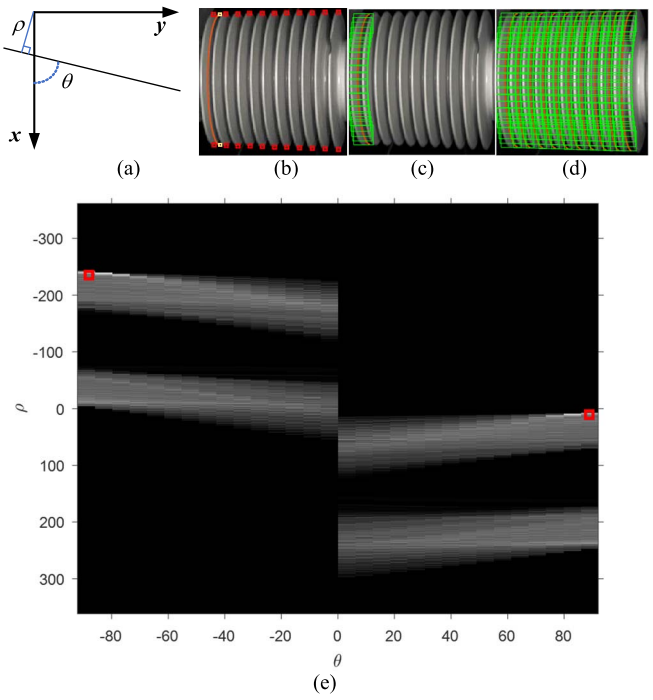


Fig. 8. Extract local insulators by the HT&PE. (a) Definition of θ in (d). (b) Detected collinear points and ellipse-sliding path. (c) Window sliding along a path. (d) Local insulator patches obtained through the insulator. (e) Hough matrix.

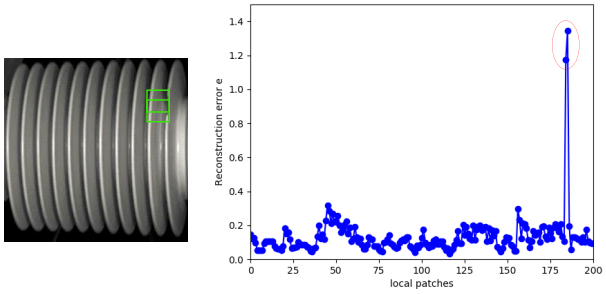


Fig. 9. Reconstruction errors for patches in an insulator image.

B. Defect Detection by a GAN-2 Model

The structure of the GAN-2 model is similar to that of the GAN-1 model, which is shown in Fig. 6. However, the GAN-2 model is used for distinguishing insulator states, while the GAN-1 model is used for localization refinement. We utilize normal insulator patches rather than global insulators to train the GAN-2 model because small defects on the global insulator are not obvious and cannot be detected correctly. The states of the insulator can be distinguished according to the reconstruction error *e,* which is expressed in formula (6). Fig. 9 shows the e of patches in an insulator, and the e of the defective patch is much larger than the e of a normal patch.

Based on the reconstruction error, we define an adaptive *Defect_score*, which is normalized to [0, 1], to quantitatively assess the defect of the insulator patch.

$$Defect\_score = \begin{cases} \dfrac{e_i - median(e_1, \ldots, e_i, \ldots, e_N)}{e_{\max} - e_{\min}}, \\ \quad if\ \ e_i \geq median(e_1, \ldots, e_i, \ldots, e_N) \\ 0, \quad if\ \ e_i < median(e_1, \ldots, e_i, \ldots, e_N) \end{cases} \tag{8}$$

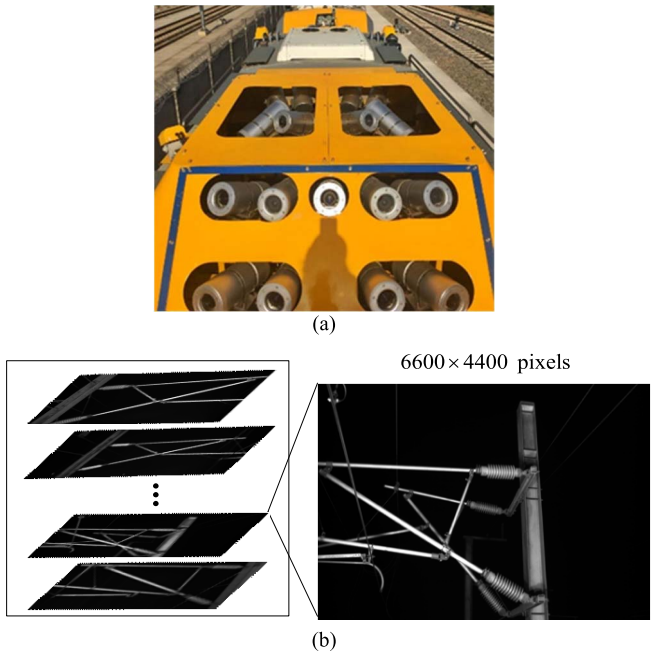


Fig. 10. Image data collection. (a) JX-300 inspection vehicle. (b) Collected catenary images.

where N is the number of insulator patches in an insulator image and $e_{\max}$ and $e_{\min}$ are the maximum and minimum reconstruction errors in an insulator image, respectively. Considering that the defect region in an insulator image is always smaller than the normal region, the *Defect_score* of the insulator patch is 0 when its e is less than the median error. When the e of a patch is much larger than the median error, the *Defect_score* of this patch is high. In this article, we predefined a threshold T_d for defect detection decisions, and an insulator patch is considered defective when its *Defect_score* exceeds the threshold T_d

V. Experiments and Analysis

A. Dataset and Platform

The image dataset adopted in this study are 6600 × 4400 pixels images that were collected by the JX-300 inspection vehicle, as shown in Fig. 10. The used dataset was collected from a 60km section in the railway line from Changsha to Zhuzhou in China. It has in total 2,479 catenary images, which include approximately 7,450 insulators. We use 1,240 catenary images to train the TOL-Framework and 1,239 images for testing. For GAN training, 3,000 standard insulator images are used to train the GAN-1 model, and 100,000 insulator patches are used to train the GAN-2 model. To build a dataset that can be directly used for deep learning models for oriented localization, we apply a special oriented annotation tool, Collabeler, to annotate each insulator in both the training and testing sets, the latter of which is just required for evaluation. In each annotation, the insulator is manually assigned an oriented rectangular bounding box and a category tag.

The experimental environment of all the implemented deep learning models is as follows: TensorFlow deep learning

framework, Linux Ubuntu 14.04, Intel Xeon CPU E3-1230 V2 clocked at 3.3 GHz, 12 GB RAM, and GTX1080Ti GPU with 11 GB memory.

B. Performance Evaluation of Localization

1) Training Parameters and Evaluation Criteria: For fair comparisons, all the deep learning models are trained by using similar parameter settings, which are as follows. The momentum and weight decay are set to 0.9 and 0.0005, respectively. The total number of iterations is set to 40,000. The learning rate is initialized a large number (0.001) and then decayed at the 20,000$^{\text{th}}$ and 30,000$^{\text{th}}$ iterations by multiplying by 0.1. For the evaluation criteria, the widely used precision, recall and F1-score are adopted, which are expressed in formulas (9)-(11), respectively, where *TP* is the number of correctly classified insulators, *FP* is the number of objects that are misclassified as insulators, and *FN* is the number of insulators that are misclassified as other objects. In particular, the AP (average precision) is adopted to evaluate the overall localization performance, as expressed in formula (12). The AP is equal to the area under the precision-recall (P-R) curve.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{9}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{10}$$

$$\text{F1-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

$$\text{AP} = \int P(R) dR \tag{12}$$

2) Verification of the Effectiveness of the Proposed Modules: To verify the localization improvements produced by the selected FPN and the proposed modules (the cascade regression and post-refinement modules), we take the oriented Faster R-CNN ResNet-101 model as the baseline and progressively add the proposed modules to the baseline to construct different models for comparison. All of these models are trained and tested on our built catenary dataset. Note that, unlike the existing methods that use a loose IoU threshold (i.e., 0.5 or 0.7) to evaluate the localization performance, we also apply strict IoU thresholds. In this article, we set the IoU threshold to 0.5, 0.8 and 0.9 and $|\Delta\varphi| < 10^\circ$ (the angle difference between the ground-truth bounding box and predicted bounding box is less than 10°) to give a more comprehensive evaluation. The obtained localization performances of the compared models are shown in Tabs. I-III, and the corresponding P-R curves are displayed in Figs. 11(a)-(c). We compare and analyze these results as follows.

1) The models with FPN, cascade regression or post-refinement modules achieve higher precision, recall, F1-score and AP than the baseline model, especially when a high IoU threshold is applied. As shown in Tab. III and Fig. 11(c), the AP of the baseline model drops greatly from 0.914 to 0.159 when the IoU threshold is changed from 0.5 to 0.9, but the AP of the proposed TOL-Framework is 0.804 when the IoU threshold is 0.9. The AP improvement brought by FPN, cascade regression or post-refinement

TABLE I

LOCALIZATION PERFORMANCES WITH AN IOU THRESHOLD OF 0.5

Oriented Faster R-CNN ResNet-101 (baseline)			Recall	Precision	F1-score	AP
FPN	Cascade regression module	Post-refinement module				
			0.917	0.972	0.943	0.914
√			0.993	0.985	0.989	0.981
√	√		0.999	1.000	0.999	0.999
√	√	√	--	--	--	--

TABLE II

LOCALIZATION PERFORMANCES WITH AN IOU THRESHOLD OF 0.8

Oriented Faster R-CNN ResNet-101 (baseline)			Recall	Precision	F1-score	AP
FPN	Cascade regression module	Post-refinement module				
			0.856	0.906	0.881	0.795
√			0.963	0.956	0.959	0.926
√	√		0.995	0.996	0.995	0.992
√	√	√	--	--	--	--

TABLE III

LOCALIZATION PERFORMANCES WITH AN IOU THRESHOLD OF 0.9

Oriented Faster R-CNN ResNet-101 (baseline)			Recall	Precision	F1-score	AP
FPN	Cascade regression module	Post-refinement module				
			0.377	0.399	0.388	0.159
√			0.601	0.597	0.599	0.367
√	√		0.764	0.765	0.765	0.594
√	√	√	0.894	0.895	0.895	0.804

can be observed obviously from the areas under the PR curves in Fig. 11, and the quantitative AP values are shown in Tabs. I-III. This finding indicates that our proposed modules are effective in obtaining tighter oriented localization.

2) When the IoU threshold is set to a low value (0.5), all the compared methods have high APs (over 0.9), as shown in Tab. I and Fig. 11(a). However, all methods' APs decrease with the growth of the IoU threshold value, and their performances have obvious differences, which indicates that the low IoU threshold used in the existing methods [25], [34] is not suitable for evaluation and the high IoU threshold is effective for performance evaluation.

3) The recall and precision of the baseline + FPN + Cascade regression model are near 1 when the IoU threshold is 0.8, as shown in Fig. 11(b). When the IoU threshold exceeds 0.8, the recall and precision are less than 1. Therefore, 0.8 is the extreme IoU threshold for the baseline + FPN + Cascade regression model, which can also be observed in the *AP vs IoU* curve of the TOL-Framework (Fig. 12). Thus, we applied post-refinement to the baseline + FPN + Cascade regression model when the IoU threshold is 0.8 to ensure candidate bounding box generation is implemented in a minimum region and all insulators are correctly classified.

TABLE IV
COMPARISONS OF THE OVERALL PERFORMANCE

Method	TP	FP	FN	F1-score	FPS (Frame/s)
Faster R-CNN with an AE [25]	53	16	29	0.702	0.52
Faster R-CNN with a GAN model [34]	36	27	46	0.497	0.81
TOL-Framework with GAN-1 (global insulator)	40	22	42	0.556	0.32
TOL-Framework with GAN-2 (insulator patch)	79	6	3	0.946	0.27

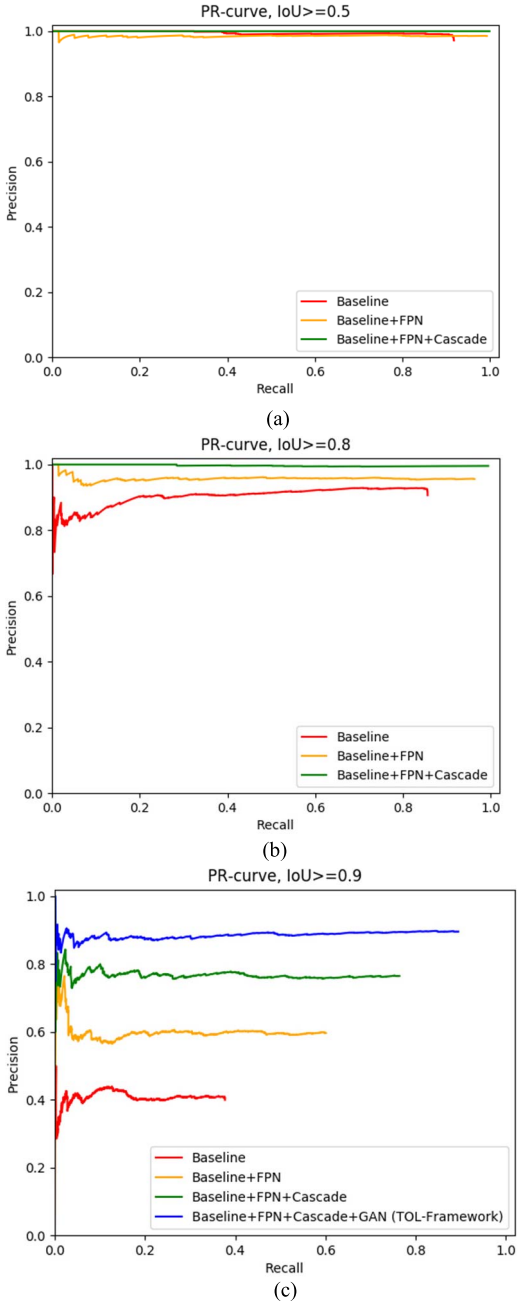


Fig. 11. PR curves of the different localization models. (a) PR curve with an IoU threshold of 0.5. (b) PR curve with an IoU threshold 0.8. (c) PR curve with an IoU threshold 0.9.

The above presents a quantitative analysis of the localization performance. The following Figs. 12(a)-(c) also display some localization results produced by different methods for comparison, which shows that the proposed modules can help to obtain tighter oriented boxes and achieve high localization accuracy than the comparison methods. The size of the trained TOL-Framework is 1,324 Mb, and the test efficiency is 0.332 FPS (frame per second). It can be run successfully on the GTX1080Ti GPU.

3) Comparison of the TOL-Framework With Other Localization Methods: We compare the proposed TOL-Framework with other oriented localization methods, namely, R2CNN [20], RRPN [21], R-DFPN [22] and the widely used horizontal localization method Faster R-CNN [13]. To analyze the post-refinement model, the TOL-Framework without post-refinement is used for comparison. The localization performances of all compared methods are shown in Fig. 13, which are the AP vs IoU threshold curves.

1) When the IoU threshold is less than 0.8, only our proposed TOL-Framework and the TOL-Framework without post-refinement have the highest APs, which are close to 1. According to formulas (6), (7) and (9), both the recall and precision are equal to 1, which indicates that the proposed methods can localize all the insulators, but the localized insulators may have low overlaps with their true positions. Moreover, the APs of the other compared methods are less than those of our methods, which indicates that some insulators cannot be localized by these comparison methods even when the IoU threshold is set to a low value. 2) When the IoU threshold is larger than 0.8, the APs of our methods decline slower than those of other methods, which indicates that the insulator boxes produced by our methods are closer to their true positions. 3) As the green curve in Fig. 13 shows, we apply the post-refinement module to baseline + FPN + Cascade regression when the IoU threshold is set to 0.8 because some insulators cannot be localized (AP $<$ 1) when the IoU threshold is larger than 0.8 and inaccurate localization (AP $\approx$ 1, but overlaps with true positions are small) will lead to time cost increases and performance degradation in the post-refinement module. 4) The TOL-Framework achieves higher APs than the horizontal localization Faster R-CNN model for all IoU thresholds. The comparison results show the proposed TOL-Framework performs better than other localization methods based on neural networks.

C. Overall Performance Evaluation for Defect Detection

In our experiment, the threshold T_d of *defect_score* for defect detection is set to 0.65, and the GAN-2 model is trained for 50 epochs. The learning rate is initialized at 0.0002, and it is decayed by multiplying it by 0.5 at the 40th epoch. Our dataset includes 82 defect insulators, all of which can be used

(a) (b) (c)

Fig. 12. Localization cases produced by different methods. (a) Localization of the baseline + FPN model. (b) Localization of the baseline + FPN + Cascade regression model. (c) Localization of the baseline + FPN + cascade regression + post-refinement (TOL-Framework) model.

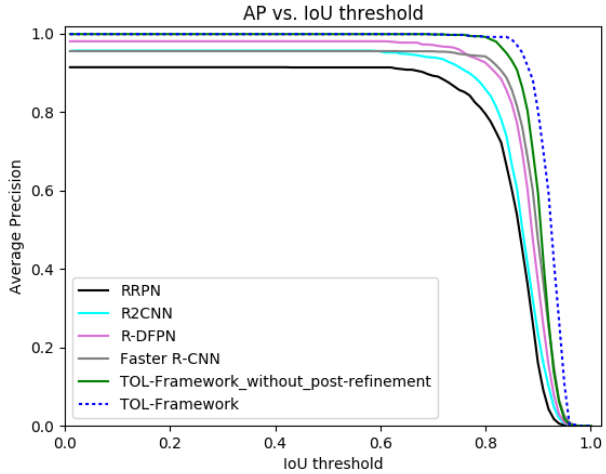


Fig. 13. Localization performances of all the compared methods.

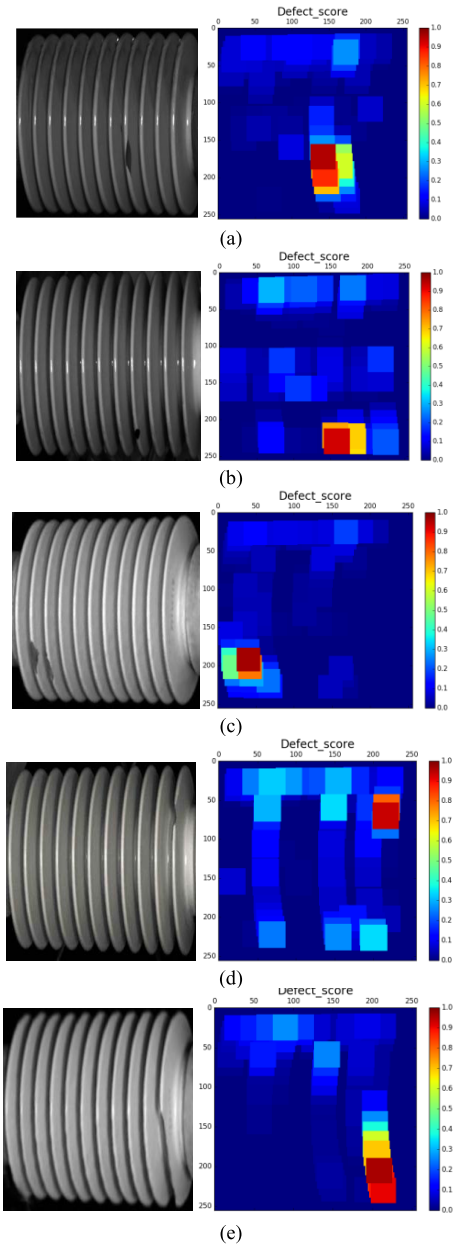


Fig. 14. Defect_score maps of different defective insulators produced by the proposed GAN-2 model.

for testing because the proposed GAN-based defect detection method does not need defective samples for training. During testing, 79 defect insulators are correctly detected, and 3 defect insulators are not detected by using *the TOL-Framework with GAN-2*, which are reported in *TP*s (true positives) and *FN*s (false negatives), respectively, in Tab. IV. In Fig. 14, examples of defect detection results for different insulators in the China railway line are presented. The *Defect_score* maps of defective patches show higher values in the locations with defects. In particular, small defects are correctly detected, as shown in Fig. 14(b) and Fig. 14(d). In addition, the proposed method can find the position of the defect on the insulator. The obtained detection results can be directly used as maintenance decision support for asset managers.

We compare not only the performances of two different GANs based on the global insulator and insulator patches but also our method with the current catenary insulator detection methods [25], [34]. The quantitative results are presented in Table IV. It shows the proposed method "TOL-Framework with GAN-2" has the highest accuracy (F1-score) 0.946. The speed of the proposed method is 0.27 (Frame/s), which is slower than other methods. As the current catenary detection system is implemented off-line, accuracy is the most important feature. The obtained results can be provided to asset managers as maintenance decision support information.

In this article, the proposed method is evaluated with an image dataset collected from a section of the rail line Changsha-Zhuzhou. The method can be directly applied to the rest of this rail line as well. Note that there are also other railway lines in China with the same configuration and type of catenary system. This means that the proposed method

can be applied to these railway lines as well using the same inspection vehicle for image collection. For other railway lines with a different type of insulator or a different configuration of catenary, a new dataset from the new railway line should be built and the model parameters should be tuned accordingly.

VI. Conclusions

This article proposed a novel method for catenary defect detection. A new localization method, the TOL-Framework, is applied to realize tighter oriented localization, which can greatly reduce the harmful effects of incomplete localization and unnecessary background information and provide suitable insulator images for defect detection. For defect detection, a GAN-based method is proposed to evaluate the errors between normal samples and defective samples to realize defect detection. The GAN-based method can solve the problem of insufficient available defective simples with supervised learning. Experiments show that our method is effective for insulator defect detection.

There are still several aspects that need to be further considered:

1) Generalization of the method, including analysis of other railway lines with different catenary systems.
2) The efficiency of the post-refinement network can be further improved. Developing an adaptive algorithm for the selection of candidate boxes may help to reduce the time cost of our method.
3) The camera captures 2D images, which may not include all sides of the insulator. Thus, 3D inspection techniques could be attempted in the future.

References

[1] S. B. Gao, Z. G. Liu, and L. Yu, "Detection and monitoring system of the pantograph-catenary in high-speed railway (6C)," in *Proc. 7th Int. Conf. Power Electron. Syst. Appl. - Smart Mobility, Power Transf. Secur. (PESA)*, Hongkong, Dec. 2017, pp. 779–788.

[2] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[3] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, San Diego, CA, USA, Jun. 2005, pp. 886–893.

[4] X. Wang, T. X. Han, and S. Yan, "A hog-lbp human detector with partial occlusion handling," *Proc. IEEE Int. Conf. Comput. Vis.*, Kyoto, Japan, Sep. 2009, pp. 32–39.

[5] J. Zhong, Z. Liu, G. Zhang, and Z. Han, "Condition detection of swivel clevis pins in overhead contact system of high-speed railway," *J. China Railway Soc.*, vol. 39, no. 6, pp. 65–71, Jun. 2017.

[6] 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.

[7] Y. Han, Z. Liu, X. Geng, and J. P. Zhong, "Fracture detection of ear pieces in catenary support devices of high-speed railway based on HOG eatures and two-dimensional Gabor transform," *J. China Railway Soc.*, vol. 39, no. 2, pp. 52–57, 2017.

[8] H. Fan, P. C. Cosman, Y. Hou, and B. Li, "High-speed railway fastener detection based on a line local binary pattern," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 788–792, Jun. 2018.

[9] H. Azizpour and I. Laptev, "Object detection using strongly-supervised deformable part models," in *Proc. Eur. Conf. Comput. Vis.*, Firenze, Italy, Oct. 2012, pp. 836–849.

[10] M. Pedersoli, J. Gonzalez, X. Hu, and X. Roca, "Toward real-time pedestrian detection based on a deformable template model," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 355–364, Feb. 2014.

[11] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, San Diego, CA, USA, Jun. 2014, pp. 580–587.

[12] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 779–788.

[13] 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.

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[15] J. F. Dai, Y. Li, K. M. He, and J. Sun, "R-FCN: Object detection via region-based fully convolution networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 379–387.

[16] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," 2016, *arXiv:1612.03144*. [Online]. Available: http://arxiv.org/abs/1612.03144

[17] Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving into high quality object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6154–6162.

[18] W. Liu *et al.*, "SSD: Single shot multi-box detector," in *Proc. Eur. Conf. Comput. Vis.*, Amsterdam, The Netherlands, 2016, pp. 21–37.

[19] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6517–6525.

[20] Y. Jiang *et al.*, "R2CNN: Rotational region CNN for orientation robust scene text detection," 2017, *arXiv:1706.09579*. [Online]. Available: http://arxiv.org/abs/1706.09579

[21] J. Ma *et al.*, "Arbitrary-oriented scene text detection via rotation proposals," 2017, *arXiv:1703.01086*. [Online]. Available: http://arxiv.org/abs/1703.01086

[22] X. Yang, H. Sun, K. Fu, J. Yang, X. Sun, M. Yan, and Z. Guo, "Automatic ship detection in remote sensing images from Google earth of complex scenes based on multiscale rotation dense feature pyramid networks," *Remote Sens.*, vol. 10, no. 1, pp. 132–145, Jan. 2018.

[23] J. Chen, Z. Liu, H. Wang, A. Nunez, and Z. Han, "Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 2, pp. 257–269, Feb. 2018.

[24] Z. Liu, L. Wang, C. Li, and Z. Han, "A high-precision loose strands diagnosis approach for isoelectric line in high-speed railway," *IEEE Trans. Ind. Informat.*, vol. 14, no. 3, pp. 1067–1077, Mar. 2018.

[25] G. Kang, S. Gao, L. Yu, and D. Zhang, "Deep architecture for high-speed railway insulator surface defect detection: Denoising autoencoder with multitask learning," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 8, pp. 2679–2690, Aug. 2019.

[26] W. Liu, Z. Liu, A. Nunez, and Z. Han, "Unified deep learning architecture for the detection of all catenary support components," *IEEE Access*, vol. 8, pp. 17049–17059, 2020.

[27] 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.

[28] 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.

[29] J. Zhong, Z. Liu, 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.

[30] H. Feng, Z. Jiang, F. Xie, P. Yang, J. Shi, and L. Chen, "Automatic fastener classification and defect detection in vision-based railway inspection systems," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 4, pp. 877–888, Apr. 2014.

[31] X. Gibert, V. M. Patel, and R. Chellappa, "Deep multitask learning for railway track inspection," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 153–164, Jan. 2017.

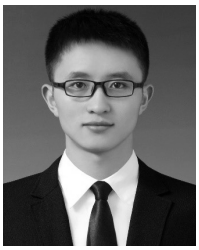
[32] G. E. Hinton, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[33] I. Goodfellow *et al.*, "Self-attention generative adversarial networks," in *Proc. Mach. Learn.*, 2014, pp. 1–5.

[34] 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.
[35] Q. Xuan, Z. Chen, Y. Liu, H. Huang, G. Bao, and D. Zhang, "Multiview generative adversarial network and its application in pearl classification," *IEEE Trans. Ind. Electron.*, vol. 66, no. 10, pp. 8244–8252, Oct. 2019.
[36] Z. X. Zhao, B. Li, R. Dong, and P. Zhao, "A surface defect detection method based on positive samples," in *Trends in Artificial Intelligence*, vol. 11013. Cham, Switzerland: Springer, 2010.
[37] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, "GANomaly: Semi-supervised anomaly detection via adversarial training," 2018, *arXiv:1805.06725*. [Online]. Available: http://arxiv.org/abs/1805.06725
[38] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.

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 system and its automation from Southwest Jiaotong University, Chengdu, China, in 2003. He is currently a Full Professor with the School of Electrical Engineering, Southwest Jiaotong University. His current research interests include the electrical relationship of EMUs and traction and detection and assessment of pantograph-catenary in high-speed railway. He is a fellow of IET. He is an Associate Editor for the journal the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT. He is also on the Editorial Board of the journal the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.



Cheng Yang (Student Member, IEEE) received the B.S. degree in power system and its automation from Southwest Jiaotong University, Leshan, China, in 2018, where he is currently pursuing the master's degree in electrical engineering. His current research interests include computer vision and their applications in railway fault detection

Hongrui Wang (Member, IEEE) received the B.S. degree in electrical engineering and automation from Mao Yisheng Class, Southwest Jiaotong University, Chengdu, China, in 2012, and the Ph.D. degree with the Section of Railway Engineering, Delft University of Technology, Delft, The Netherlands, in 2019. His research interests include signal processing, machine learning, and their applications in the condition monitoring and maintenance of railway infrastructure. He is an Associate Editor for the journal the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT.

Shibin Gao received the Ph.D. degree in electrical engineering from Southwest Jiaotong University, Chengdu, China. Since 1998, he has been a Full Professor with the Department of Electrical Engineering, Southwest Jiaotong University. His current research interests include power system protection and automation, monitoring of electrical equipment, rail transit traction power supply system security, and railway infrastructure service status monitoring.

Alfredo Núñez (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Universidad de Chile in 2010. He was a Post-Doctoral Researcher with the Delft Center for Systems and Control, Delft, The Netherlands. Since 2013, he has been with the Section of Railway Engineering, Department of Engineering Structures, Delft University of Technology. He is currently an assistant professor (tenured) in data-based maintenance for railway infrastructure. His current research interests include the maintenance of railway infrastructures, intelligent conditioning monitoring in railway systems, big data, risk analysis, and optimization. He is on the Editorial Board of the journal *Applied Soft Computing* and an Associate Editor of the journal the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS.