Automatic defect detection based on improved Faster RCNN for substation equipment

Lv Shouguo¹, Liu Kai²*, Qiao Yaohua¹, Wang Qi³, Sun Yang¹ and Li Zhenyu²

¹State Gird Shandong Maintenance Company, Jinan, SD 531, China
²State Grid Intelligent Technology Co., Ltd, Jinan, SD 531, China
³Corresponding author’s e-mail: 1997.liukai@163.com

Abstract. Defect detection methods based on machine learning extremely accelerate the substation routine inspection process. In this paper, we propose an automatic defect detection method based on improved Faster RCNN. For one thing, random feature pyramid (RFP) structure is introduced for the highly discriminative feature map construction; for another thing, we execute the detection boxes selection by soft non-maximum suppression (SNMS), keeping the detection of defects which distribute densely. Finally, online hard example mining (OHEM) is employed to deal with the imbalance problem. Experimental results demonstrate that the proposed approach obtains competitive performance compared with state-of-the-art deep learning object detection methods.

1. Introduction
Substation routine inspection, which aims to check the work station of equipment, is vitally critical for the stable operation of power transformer substation. Recently, with the development of video surveillance and robot technology, employing robot or video camera for equipment monitoring has become the new trend of substation inspection. However, robot and video camera can only deal with the simple functions, such as video/image capture, storage and playback, the most critical work, image/video analysis and defect detection, relies on human with rich experience to check the inspection images recordings visually. The visual check of huge number of images is very tedious and time consuming. Therefore, the automatic detection technology is valuable for assisting power engineer to analyse the substation inspection recordings.

The methods for automatic substation equipment defect detection have been under study for several years. Some studies search for robust appearance features for discriminating the normal equipment and defect ones. Liu et al. [1] developed an image-based state recognition approach, and they propose extraction of texture features using a Gabor transformation, then the state of the isolator is classified by the SVM. Prasad et al. [2] employed LBP-HF features to describe the appearance of substation equipment. However, the uncertain defect appearance variations make the robust features hard to design. Inspired by the good performance of deep learning methods in multiple classifications and recognition tasks recently, some studies leveraged deep learning methods to discover an optimal substation equipment defect feature representation scheme and corresponding classifier. Pan et al. [3] introduced a power line detection method based on CNN and Hough transform. Nordeng et al. [4] proposed a DEBC detection based on Fast-RCNN.

Although the methods based on deep learning-based have achieved encouraging performance in substation equipment defect detection task, they still have some drawbacks. For one thing, the equipment
defects with small size are apt to be omitted during the convolution process of deep learning. For another, the class imbalance (i.e., the number of normal equipment is much bigger than that of defective equipment) make the deep learning methods misjudge the normal equipment as defect frequently, result in relatively low precision rate.

To address such drawbacks, we propose an improved Faster RCNN method for substation equipment defect detection in this paper. Firstly, features with different resolution are selected randomly and combined together, making the Faster RCNN have strong detectability on equipment with variable size. Secondly, the Soft Non-Maximum Suppression (SNMS) process is introduced to prevent the loss of densely distributing defect. Thirdly, Online Hard Example Mining (OHEM) is adopted to train the Convolutional Neural Networks (CNN) for classify normal equipment and defective equipment. The performances of the different approaches are compared in terms of recall, precision and average precision (AP).

The remainder of the paper is organized as follows. Section 2 briefly describes the substation equipment dataset used in our study. Section 3 is devoted to the proposed method, which consists of Random Feature Pyramid, soft-NMS and Online Hard Example Mining. The experimental results and discussions are presented in Section 4 and Section 5. Section 6 concludes this paper.

2. Dataset
The substation equipment dataset come from State Grid Intelligent Technology Co., Ltd, this dataset has been widely employed in substation equipment default detection research area. The complete dataset contains 10361 images of normal equipment and defective equipment captured from three 220KV substations. In our study, 8000 images, which contain three types of defects (i.e., metal corrosion, smoking and meter glass crack) are selected randomly for algorithm training and performance evaluation. And these images are divided into 7200 images and 800 images for algorithm training and performance respectively. Figure 1 illustrates some equipment images with defects.

![Figure 1. Images of substation equipment with defects.](image)

3. Method

3.1. Faster RCNN
As shown in Figure 2, the whole process of Faster-RCNN is composed of three stages [5]. In the first stage, the substation equipment images are sent into a pretrained CNN network and the corresponding feature map is obtained. In our study, the Resnet-101 network, which contains convolutional layers, ReLU and pooling layers, is employed. In the second stage, the ROI proposal regions are generated by a region proposal network (RPN) based on the feature map. In the third stage, the selected proposal from feature map are inserted into fully-connected network and softmax network, and the final object region and its classification results are combined as the outputs of Faster-RCNN.

![Figure 2. The framework of Faster-RCNN.](image)
3.2. Random Feature Pyramid

Feature pyramid (FP) is an effective method in object detection to deal with the objects at vastly different scales [6]. By combining low-resolution, semantically strong features with high-resolution, semantically weak features via a top-down architecture, feature pyramid collects rich semantics at all levels. The architecture of feature pyramid is illustrated in Figure 3(a). However, the stubborn structure of pyramid implies lots of redundancy, weakening the flexibility of network. In this paper, we propose a Random Feature Pyramid (RFP), in which half part of the feature maps are selected randomly to build the final feature map. The architecture of random feature pyramid is shown in Figure 3(b).

3.3. Soft Non-Maximum Suppression

Non-Maximum Suppression (NMS) locates behind the CNN during object detection process. After getting a list of detection boxes B with scores S, the detection box with the maximum score M_S is selected and appended into the final detection set D. then every box in B, which has an overlap greater than a threshold N with M, is removed. A major issue with non-maximum suppression is that the score of neighbouring detections are set to zero directly. Thus, if an object was actually present in that overlap threshold, it would be missed and this would lead to a drop in average precision.

To solve this problem, we introduce Soft Non-Maximum Suppression (SNMS), in which a Gaussian function is leveraged as the scaling factor for detection scores. The formula of Gaussian function is shown as follow.

\[
f(IOU(M, b_i)) = e^{-\frac{IOU(M, b_i)^2}{\sigma}}, (\forall b_i \notin D)
\]

Where \(\sigma = 0.6\)

As can be seen, the bounding boxes, which have high overlap with M, are assigned low scores, and the bounding boxes with few overlaps get high scores.

3.4. Online hard example mining

In the practical substation routine inspection, most of the equipment are normal and defect is scarce, leading a serious class imbalance problem for defect detection. This imbalance may be as extreme as 1,000 normal equipment to 1 defect equipment. The most popular strategy to deal with the class imbalance problem is bootstrapping, which employ an iterative training algorithm that alternates between updating the detection model given the current set of examples, and then using the updated model to find new false positives to add to the bootstrapped training set.

In this paper, we introduce the online hard example mining (OHEM) [7] to deal with the imbalance between normal equipment and defect equipment. During the iterative train process, substation equipment images are sampled according to a non-uniform, non-stationary distribution that depends on the current loss of each images under consideration, thus in each iteration, only the samples with high loss construct the training batch.

4. Results

All the experiments were carried out in PyTorch deep learning framework based on Ubuntu16.04 operating system. The GPU card is NVIDIA GTX1080. For three type of typical substation equipment defects, the performance of the algorithm can be assessed by the statistical measures of recall, precision and Average Precision.

In our study, traditional feature pyramid strategy and random feature pyramid strategy are separately employed for feature map construction. The average results for three defects are listed in Table 1. It can be noticed that the recognition accuracies are almost same for metal corrosion and meter glass crack.
However, the random feature pyramid makes the recall and precision of smoking higher than traditional feature pyramid.

| Defect type         | Recall  | Precision |
|---------------------|---------|-----------|
| metal corrosion     | 0.636   | 0.635     |
| smoking             | 0.845   | 0.879     |
| meter glass crack   | 0.863   | 0.869     |

In Table 2, the performance of Faster RCNN with NMS and soft NMS are compared. It can be noticed from the table that after employing S-NMS, the recall rate of metal corrosion increased by ~14% as compared with that of common NMS.

Further, the detection performances for these defects are compared with No-OHEM and OHEM training strategies. The recall and precision rate under these two experimental conditions respectively are shown in Table 3.

| Defect type         | Recall  | Precision |
|---------------------|---------|-----------|
| metal corrosion     | 0.636   | 0.729     |
| smoking             | 0.845   | 0.903     |
| meter glass crack   | 0.863   | 0.925     |

Table 4 presents a comparison on AP between the proposed methods and YOLOv3, SSD and Faster-RCNN. The results in the table show that our method improves recognition accuracy of metal corrosion, smoking and meter glass crack, indicating the better performance of our method. Representative detection results are shown in Figure 4.

| Method              | metal corrosion | smoking | meter glass crack |
|---------------------|-----------------|---------|-------------------|
| YOLOv3              | 0.632           | 0.841   | 0.876             |
| SSD                 | 0.618           | 0.819   | 0.865             |
| Faster-RCNN         | 0.629           | 0.836   | 0.872             |
| The proposed method | 0.648           | 0.853   | 0.878             |

5. Discussion
From a visual perspective, the substation equipment defects with small or variational size, such as metal corrosion, are the biggest challenge for automatic defects detection work. The reason of CNN based
The method’s failure is the multiple convolutions, which damage the visual information of small defective objects.

RFP strategy collects half part of the feature maps randomly for the final feature map and overcomes such problem, which can be seen in Table 1. Moreover, the traditional NMS method, which set the score of neighbouring detections to zero, is cruel for the defect that distribute densely. The proposed SNMS employ a Gaussian score function to save the adjacent object around the candidate defect. It can be found from Table 2 that the recall and precision of metal corrosion with SNMS are generally higher than that of NMS. The equipment with defect only take up very few parts of total substation equipment, resulting in severe negative-positive imbalance problem. Bootstrapping is the common method to deal with such problem. OHEM, which pays more attention to the samples with high loss in iterative training process, can improve the recall and precision of all defects generally. The experiment results in Table 3 confirm the effect of OHEM. Finally, As can be seen from the Table 4, the proposed method outperforms other CNN-based substation equipment defect detection methods.

Although our work focuses on the detection of equipment defects, future research is under way to devise an online automatic defect detection system for power supply company based on the proposed method. Firstly, the video and images captured by camera and robot are sent to database system. Secondly, a video-frame drawing technique is employed to divide surveillance video into numbers of frames. Thirdly, the detection method proposed in this paper is carried out for each image, to determine whether any defects are shown in image and the location of defects. The OHEM technique we used in this paper can deal with the heavily class-imbalanced problem, i.e., defect equipment are extremely rare compared to normal ones. In many cases, equipment images are recorded with several camera or robot in different sites of the substation. If a detect detection system can take the classification results from all images into account and make the decision with a predefined criterion, the recall rate and precision rate are likely to increase due to the integration of information.

6. Conclusions
In this paper, we propose an automatic defect detection method based on improved Faster RCNN. Firstly, random feature pyramid (RFP) structure is introduced for the highly discriminative feature map construction. Secondly, we execute the detection boxes selection by soft non-maximum suppression (SNMS), keep the detection of defects which distribute densely. Finally, online hard example mining (OHEM) is employed to deal with the imbalance problem. The recall, precision and average precision of different methods are compared and the proposed approach obtains competitive performance compared with other deep learning-based object detection methods.

Acknowledgments
This work was supported by the technology project of State Gird Corporation. The specific project name is "Research and application of key technologies of transmission line UAV intelligent patrol based on artificial intelligence ". Thanks very much.

References
[1] Liu S., Xu X. J. Substation Intelligent Monitoring System Based on Pattern Recognition. In: International conference on applied informatics and communication, Xi An. 2011. pp. 473-478.
[2] Prasad P. S., Rao B. P. LBP-HF features and machine learning applied for automated monitoring of insulators for overhead power distribution lines. In: International Conference on Wireless Communications, Signal Processing and Networking. Chennai, India. 2016. pp. 808-812.
[3] Pan C., Cao X., Wu D. Power line detection via background noise removal. In: IEEE Global Conference on Signal and Information Processing. Washington, DC, USA. 2017. pp.871-875.
[4] Nordeng I. E., Hasan A., Olsen D., et al. DEBC detection with deep learning. In: 20th Scandinavian Conference on Image Analysis(SCIA). Cham, Springer. 2017. pp.248-259.
[5] Ren S., He K., Girshick R., et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. (2017) IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6):1137-1149.

[6] Lin T. S., Piotr Dollar, Girshick R., et al. Feature Pyramid Networks for Object Detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society. Hawaii, USA. 2017. pp:2117-2125.

[7] Shrivastava, Abhinav, Gupta, Abhinav, Girshick, Ross. Training Region-Based Object Detectors with Online Hard Example. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, Las Vegas, NV, USA. 2016. pp:761-769.