An accurate and fast detection method for railway lightning protection component

Suining Wu*, Bin Li¹, Zhixin Wang¹, Xiaoguang Chen¹, Long Shi ¹ and Shulin Tan¹

¹ Institute of Intellectual Technology, China Railway Signal & Communication Research & Design Institute Group Co., Ltd., Beijing, 100070, China

*Corresponding author’s e-mail: wusuining@crscd.com.cn

Abstract. Lightning protection component is used to display the working condition of railway cable networks. Lightning strike damage cable networks, hence, the research of detection method for railway lightning protection is of great importance. So far, there are limited methods to solve this problem. Existing methods are based on electrical and magnetic signals. However, these methods can not meet the acquirement of speed and accuracy. In this paper, we proposed a accurate and fast detection method for railway lightning protection component. The proposed method is based on deep neural network named W-YOLO. To solve the imbalanced samples problem of fault components and common components, we devised a novel classification loss function according to the proportion components. Since the raw class and location coordinate of components are lack of meaning, a sorting method of elements based on the angle and distance threshold is proposed. From the experiments, the proposed detection method is able to detect railway lightning protection component accurately in real time.

1. Introduction

Lighting strike leads greatly danger for railway cable network, which is harm for the safety of passengers and freight. At present, non-visual methods are mainly used to detect lightning protection components, such as electrical and magnetic signals[1-2]. Comparing with traditional manual detection method and non-vision detection methods, vision detection methods have the advantage of non-contact, high accuracy, dynamic and so on.

With the development of deep learning, lots of modern object detection methods based on convolutional neural networks(CNN) are proposed. Compared with traditional image process methods, they make detection accuracy and speed improved a lot. Thereby, we proposed an accurate and fast detection method for railway lightning protection component based on CNN.

In fact, the most useful results are the index of component other than the raw coordinates. Hence, we divide the detection task into two sub-task: component detection and component sort. Based on the analysis above, the following works are accomplished in this paper. First, we devised a detection framework based on CNN. To improve the detection ability for fault components, a novel classification loss function is proposed. In the last, we proposed a sort method to determine the unique index of component.

The rest of this paper is organized as follows. Section 2 introduces the related works about object detection. Section 3 introduces the proposed detection method for railway lightning protection component in detail. In section 4, experiments are conducted to illustrate the accuracy and feasibility of the proposed method. Conclusions are drawn in section 5.
2. Related works

2.1. Detection method based on low-level image pattern

Before the largely application of CNN, classic artificial designed features are most popular to realize object detection, such as Haar-like features, histogram of oriented gradient features and so on. Viola-Jones\cite{3} proposed a face detection method based on Harr feature. HOG based methods are firstly proposed in pedestrian or general object detection, such as the HOG\cite{4} and DPM\cite{5}. These methods are limited in assigned tasks, that is, the performance drops when they are adopted in other tasks. What’s more, they work very slow. Further works are conducted to improve the above methods. However, they are still not comparable to methods based on CNN.

2.2. Detection method based on CNN

Related works of CNN for object detection can be divided into two kinds: one-stage methods and two-stage methods. Two-stage methods inference the proposal boxes at first, then the following networks inference the classification by the pooled feature map. R-CNN series\cite{6-8} are the most successful two-stage object detection networks. Thereafter, lots of methods further improve the speed and precision of R-CNN series. However, the speed and parameter quantity cannot meet the firm requirement in industry.

Unlike two-stage methods, one-stage methods do not rely on proposal regions. Detection time of one-stage methods is significantly less than two-stage methods, meanwhile, the accuracy of one-stage methods is sufficient for practical application. Naturally, one-stage methods are widely used in industry, especially automatic driving, face recognition and so on.

SSD\cite{9-10} and YOLO families\cite{11-13} are most successful one-stage object detectors. They are able to realize the fast and accurate object detection at the same time. YOLOv3 is the latest and best method combining the FPN and multi scale detection.

3. Algorithm

The end-to-end YOLO detectors avoid using region proposal methods. YOLO detectors utilized raw images as the input and output the bounding boxes and classification of object. The detection time of YOLO detectors is significantly less than two-stage detectors. YOLO detectors perform fast and accurate at the same time. Moreover, several tiny-YOLO networks are proposed to further decrease parameters and improve speed.

Previous versions of YOLO predict on the highest level feature maps, which make them can achieve relatively high detection accuracy. However, localization errors occurs when YOLO performs classification and localization simultaneously. YOLOv1 and YOLOv2 inference the class and bounding boxes via the highest feature maps. However, the last convolutional layer is often spatially course, which is not sufficient for localization. Inspired by FPN\cite{14}, residual network and multi-feature map detection are used in YOLOv3 to improve the detection accuracy. Especially, the detection accuracy for small and medium size objects is improved by YOLOv3.

Another challenge to improve YOLO detectors is foreground-background class imbalance at training. He et al.\cite{15} devised focal loss to solve foreground-background class imbalance problem. With this function, detection network focus on more difficult negative examples. However, Redmon report that focal loss cannot improve YOLOv3. Derakhshani et al.\cite{16} proposed assisted excitation method to help the detection network train better with the manually excited activations.

3.1. Algorithm

The proposed detection framework is illustrated in figure 1. W-YOLO utilize YOLOv3 as the backbone. As shown in figure 1, W-YOLO simultaneously inference the object class and location according to three different scale feature maps. Thereafter, W-YOLO achieves highly precision in object detection. YOLOv1 and YOLOv2 are limited to detect small size object. Combining the upsampled feature maps from high level, the semantic information is improved. As a results, W-YOLO is able to improve the detection for railway lightning protection component. Non maximum suppression is used to delete
duplicated and inaccurate results. At last, the accurate bounding box coordinates are used to determine the index of each component.

![Detection framework of proposed method](image)

**Figure 1. Detection framework of proposed method**

### 3.2. Loss function

The original total loss of YOLOv3 can be divided into: classification loss, location loss and confidence loss. In W-YOLO, location loss and confidence loss remain the same.

We find that failure components are extremely less than common components in real images. If we directly train the network with raw dataset, the network will tend to detect an object as failure. In practical, the most important demand of our network is to detect failure components accurately. In the original loss function of YOLOv3, classification losses of two kinds of components contribute equally as shown in equation (1).

\[
L_{\text{class}} = L_{\text{class}}^{N} + L_{\text{class}}^{F}
\]  

(1)

In this paper, inverse proportion of the quantity is utilized as the weight to adjust the loss dynamically. Thereafter, the network will learn to predict two kinds of components more equitable. The modified loss is established in equation (2).

\[
L_{\text{class}} = \frac{n_{F}}{n_{N}} L_{\text{class}}^{N} + \frac{n_{N}}{n_{F}} L_{\text{class}}^{F}
\]  

(2)

### 3.3. Sort method

After the objects are detected, the bounding box coordinates are used to sort all the components. Figure 2 shows sorting process of lighting protection components. Firstly, we calculate the center point coordinates of each bounding box. After sorting process is completed in row and column, the unique index of lighting protection element is determined.

![Sorting process of lighting protection components](image)

**Figure 2. Sorting process of lighting protection components.**

First, the X coordinates of center points are sorted in ascending order. The difference value of coordinates are calculated. Two adjacent points define a line. For the adjacent lines, the difference value of angels are calculated. In this paper, components will be located to the next column index while both the difference value of coordinates and angels larger than threshold.
Similarly, the row index can be determined by the Y coordinates of center points. Hence, we can obtain the unique index of each component.

4. Experiments
To evaluate the efficiency of proposed method, we collect a LPC dataset which containing 300 images of lighting protection components. During the training process, data augmentation is adopted. We train the proposed detector with 8 GPUs for 50 epochs. The initial learning rate is 0.001, and it is decreased by 0.1 after 25 epochs.

Figure 3 shows the detect results via our method. Figure 3(a) and figure 3(b) are origin images captured by camera, figure 3(c) and figure 3(d) are rotated images. Results show that the proposed method is able to process rotated images and origin images accurately. Thereafter, the proposed method is accurate and stable in practical application.

![Detect results via proposed method. (a) and (b) are origin images; (c) and (d) are rotated images.](image)

In table 1, we compare W-YOLO with YOLOv3 on LPC dataset. After the modification of classification loss, mean average precision(mAP) is increased by 2.7%.

| Method    | mAP  |
|-----------|------|
| YOLOv3    | 89.4 |
| W-YOLO    | 92.1 |

Detail AP comparison for each class is established in table 2. From the table, AP of each class is increased via the proposed method. What’s more, AP difference between two class is decrease. That is, W-YOLO is able to solve label uneven problem.
Note that, we only modify the classification loss function to address the class label imbalance problem. The network architecture and detection process are not changed. Thereafter, the detection time is same with YOLOv3. For the proposed method, the detection time and sort time are 23ms and 2ms, respectively.

Figure 4 shows the detection results via our method and YOLOv3. Figure 4(a) and figure 4(b) are detection results via our method, figure 4(c) and figure 4(d) are detection results via YOLOv3. Missed detection problem happened in Figure 4(b) and figure 4(d). What’s more, a failure component is detected to be common via YOLOv3. In practical, these mistakes make serious influence. It is worth noting that sort results are accurate though errors occur when YOLOv3 is adopted.

5. Conclusions
In this paper, we proposed a detection method to monitor the state of lighting protection components. We utilize a deep CNN to detect lighting protection components. YOLOv3 is adopted as the backbone, meanwhile, a modified classification loss based on the quantity of objects is proposed to address label imbalance problem. To evaluate the proposed method, we construct a dataset of lighting protection components. From the experiments, the accuracy and efficiency perform well in application.

References
[1] De Bruin, T., Verbert, K., Babuska, R. (2016) Railway Track Circuit Fault Diagnosis Using Recurrent Neural Networks. J. IEEE Transactions on Neural Networks and Learning Systems., 1-11.

Table 2. AP Comparison between each class

| Method  | Normal | Failure |
|---------|--------|---------|
| YOLOv3  | 89.9   | 88.9    |
| W-YOLO  | 92.5   | 91.7    |
[2] Verbert, K., Schutter, B. D., Babuska, R. (2015) Exploiting spatial and temporal dependencies to enhance fault diagnosis: Application to railway track circuits. Proceedings of the 2015 European Control Conference. Linz. pp. 3052–3057.
[3] Viola, P., Jones, M. J. (2004) Robust real-time face detection. J. International journal of computer vision., 57(2):137-154,
[4] Dalal, N., Triggs, B. (2005) Histograms of oriented gradients for human detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition . Piscataway. pp. 886-893.
[5] Yan, J., Lei, Z., Wen, L., Li, S. (2014) The fastest deformable part model for object detection. IEEE Conference on Computer Vision and Pattern Recognition. Columbus. pp. 2497-2504.
[6] Girshick, R., Donahue, J., Darrell, T., Malik, J. (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. IEEE Conference on Computer Vision and Pattern Recognition.Columbus. pp. 580-587,
[7] Girshick, R. (2015) Fast R-CNN. IEEE International Conference on Computer Vision. Santiago. pp. 1440-1448
[8] Ren, S., He, K., Girshick, R., Sun, J. (2017) Faster R-CNN: towards real-time object detection with region proposal networks. J. IEEE Transactions on Pattern Analysis and Machine Intelligence., 39(6):1137-1149.
[9] Liu, W., Anguelov, D., Erhan, D., et al. (2016) SSD: Single Shot Multibox Detector. European Conference on Computer Vision. Amsterdam. pp. 21-37.
[10] Fu, C. Y., Liu, W., Ranga, A., Tyagi, A., Berg, A. C. (2017). DSSD : Deconvolutional Single Shot Detector. https://arxiv.org/abs/1701.06659.
[11] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. IEEE Conference on Computer Vision and Pattern Recognition. Piscataway. pp. 779-788.
[12] Redmon, J., Farhadi, A. (2016) YOLO9000: Better, Faster, Stronger. IEEE Conference on Computer Vision and Pattern Recognition. Honolulu. pp. 6517-6525.
[13] Redmon, J., Farhadi, A. (2018) YOLOv3: An Incremental Improvement. https://arxiv.org/abs/1804.02767.
[14] Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S. (2017) Feature pyramid networks for object detection. IEEE Conference on Computer Vision and Pattern Recognition. Honolulu. pp. 936-944.
[15] Lin, T. Y., Goyal, P., Girshick, R., He, K., Dollár, P. (2017) Focal loss for dense object detection. J. IEEE Transactions on Pattern Analysis & Machine Intelligence,. 99: 2999-3007.
[16] Derakhshani, M. M., Masoudnia, S., Shaker, A. H., Mersa, O., Araabi, B. N. (2019) Assisted Excitation of Activations: A Learning Technique to Improve Object Detectors. https://arxiv.org/abs/1906.05388.