Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 147 (2019) 331–337

Procedia Computer Science

www.elsevier.com/locate/procedia

2018 International Conference on Identification, Information and Knowledge in the Internet of Things, IIKI 2018

RCNN-based foreign object detection for securing power transmission lines (RCNN4SPTL)

Weishan Zhang[a,*], Xia Liu[a], Jiangru Yuan[b], Liang Xu[c,d], Haoyun Sun[a], Jiehan Zhou[1], Xin Liu[a]

[a] *College of computer and Communication Engineering, China University of Petroleum,Qingdao, China*
[b] *Research Institute of Petroleum Exploration & development, CNPC, China*
[c] *College of computer and Communication Engineering, Beijing University of Science and Technology, Beijing, China*
[d] *Qingdao Deep Intelligence Information Technology Co. LTD, Qingdao, China*
[e] *University of Oulu,Oulu, Finland*

Abstract

This paper proposes a new deep learning network - RCNN4SPTL (RCNN -based Foreign Object Detection for Securing Power Transmission lines), which is suitable for detecting foreign objects on power transmission lines. The RCNN4SPTL uses RPN (Region Proposal Network) to generate aspect ratio of the region proposals to align with the size of foreign objects. The RCNN4SPTL uses an end to end training to improve its performance. Experimental results show that the RCNN4SPTL significantly improves the detection speed and recognition accuracy, compared with the original Faster RCNN.

Keywords: Pattern recognition; RCNN; Power transmission line; CNN

1. Introduction

It is essential to maintain safety of power transmission lines. Foreign objects such as flying kites, balloons and plastic films hanging on transmission lines will harm distribution of high-voltage power, and pose threats to pedestrians and vehicles under the transmission lines. Therefore, it is critical to detect foreign objects in order to remove them in time.

At present, there are two main methods for detecting foreign bodies: manual line inspection and drone-based inspection. The manual inspection has great potential security risk because power transmission lines usually pass

* Corresponding author. Tel.:86-532-86983556
E-mail address: zhangws@upc.edu.cn

10.1016/j.procs.2019.01.232

through complex geographical environments such as mountains and rivers, highways and bridges. Manual inspection also has problems with low efficiency and poor results. The drone-based inspection carries cameras to inspect high-voltage transmission lines [1]. Although the drone-based inspection is not affected by geographical environments, large amounts of human efforts are still needed to determine whether there is a foreign object on the images and videos returned by drones.

There exists research on image morphology to detect foreign objects, for example there are some methods on extracting transmission lines in images [2, 3, 4, 5, 6]. The general process of image morphology-based detection is as follows. First, it uses Gaussian filter [7], median filter [8] or bilateral filter [9] to eliminate noise; then it applies Otsu [10] (maximum variance between classes) to segment background and foreground of images; finally, it utilizes the Hough transform [11] [12] to extract power transmission lines, and recognize foreign objects. It is difficult to choose a suitable gray threshold for all the images because of big differences of geographic background and influences of various weather conditions.

Deep learning has been developed rapidly recently and advances object detection and classification to a new level. The neural network has strong adaptability to geometric transformation and illumination. It can automatically generate feature descriptions based on input images. Ren et al. presented RCNN [13], which is a pioneer in region proposal-based target detection in deep learning. For algorithms of generating regional proposals, there are algorithms e.g. selective search proposed by Li et al.[14]. A series of variants of RCNN appear: SPP Net [15], Fast RCNN [16] and Faster RCNN [17]. Faster RCNN's speed and performance is better than other networks. However, at this stage, Faster RCNN is used to detect common objects such as pedestrians and fruits, and no one tries to apply it to the detection of foreign objects. Because such objects have no fixed shape, it is difficult for Faster RCNN to extract useful features, which raises difficulty of training and recognition.

This paper proposes a new neural network model - RCNN4SPTL, which is based on Faster RCNN for facilitating foreign object recognition on power transmission lines. The RCNN4SPTL model can automatically extract various relevant features of foreign objects on power transmission lines, and then detect foreign objects. This model greatly reduces the human interferes, increases work efficiency, compared with other methods.

2. RCNN4SPTL design and implementation

2.1. The RCNN4SPTL model

Fig. 1 presents the overall view of the RCNN4SPTL model. It consists of three parts. The first part is the shared convolutional neural network part (SPTL-Net), which extracts image features to produce image feature maps; the second part is the region proposal generation network (RPN). Its input is image feature maps, and the output is the candidate areas with various sizes and proportions. The last part is the classification regression network. Its input

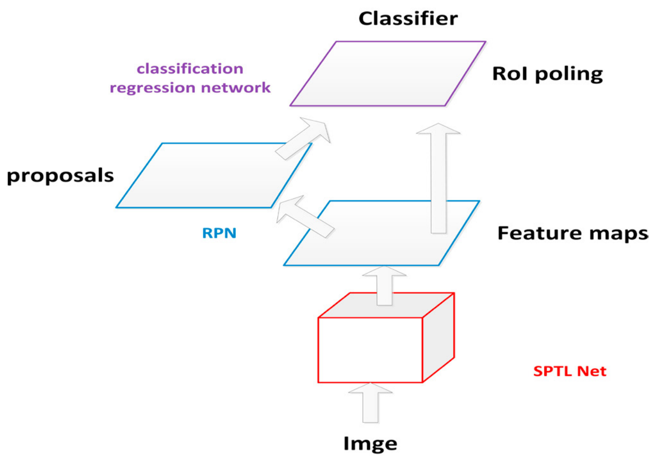

Fig. 1. The RCNN4SPTL model

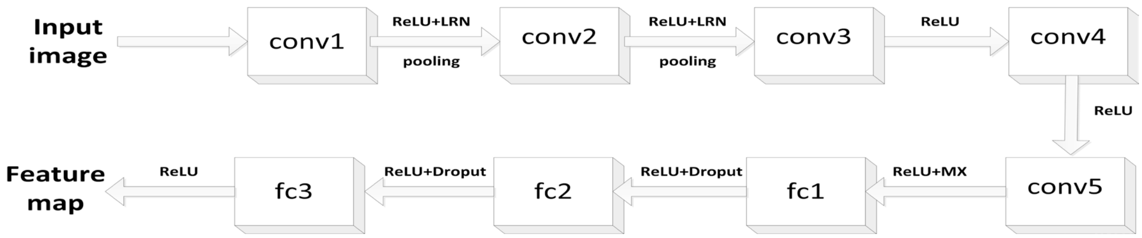

Fig. 2. SPTL-Net model

is feature maps and object region proposals. The third part produces the eigenvectors of the fixed dimension corresponding to the region proposals, and then performs image classification and localization. Finally, the RCNN4SPTL presents object categories and positions.

2.2. SPTL-Net

The RCNN4SPTL adopts SPTL-Net, which uses a smaller convolution kernel to improve the quality of feature extraction, which reduces the number of neurons to improve the training and detection speed without affecting detection performance.

The SPTL-Net is shown in Fig. 2. It has eight layers. The first five are convolutional layers, and the last three are fully connected layers. The first convolutional layer has 96 convolution kernels of size 5 x 5 x 3 filtering input images with 223 x 223 x 3. The convolution kernel has a step size of two pixels. A smaller convolution kernel is beneficial to feature fusion and subtle feature extraction. The second convolutional layer has 256 convolution kernels of size 5 x 5 x 96, convolving the pooled results from the first layer. The third convolutional layer does the same with 384 convolution kernels of size 3 x 3 x 256. The fourth and fifth convolutional layers are connected to each other without a pooling layer between them. The first fully connected layer has 4096 neurons. The number of neurons in the second fully connected layer is 1048.

The convolution and pooling operations are performed separately using equations (1) and (2).

$$output_{size} = \frac{input_{size} - kenel\_size + 2 * pad}{stride} + 1 \tag{1}$$

$$output_{size} = \frac{input_{size} - kenel\_size}{stride} + 1 \tag{2}$$

Where $output_{size}$ is the size of the output images, $input_{size}$ is the size of the input images, $kernel\_size$ is the size of the convolution kernel, pad is the size of the filled pixel, and stride is the step size.

2.3. Adjust the size and proportion of region proposals

The RPN is a convolutional neural network which uses feature maps generated by the SPTL-Net as the input and generates rectangular region proposals with various sizes and aspect ratios. The RPN first uses a 3 x 3 sliding window to slide on a feature map; it projects each position on the map passing through the window to a 256-dimensional feature vector, and then inputs each vector into next two fully connected layers. The fully connected layer with classification function produces 2 x 9 = 18 scores, and each candidate box corresponds to two scores, respectively representing the possibility of containing and not containing a given object in the candidate box. The fully connected layer with regression function produces 4 x 9 = 36 correction parameters. The RPN corrects the region proposal using these parameters, and each candidate region corresponds to four correction parameters. The anchor (the center of the current sliding window) is centered on the original image to produce region proposals with three scales and three aspect ratios. The RPN utilizes nine candidate rectangular regions to adapt to objects. The three different sizes are 128^2, 256^2, and 512^2, and the three aspect ratios are 1:1, 1:2, and 2:1.

The RPN generates four correction parameters tx, ty, tw, and th for each candidate region, and uses the four parameters to correct the region proposal. Equations (3) to (6) are the correction formulas:

$$x = w_a t_x + x_a \tag{3}$$

$$y = h_a t_y + y_a \tag{4}$$

$$w = w_a \exp t_w \tag{5}$$

$$h = h_a \exp t_h \tag{6}$$

where x, y are the x-, y-coordinates of the center point, w and h are the width, and height of the corrected candidate region. x_a and y_a indicate the horizontal and vertical coordinates of the candidate region center point, and w_a and h_a are the width and height of the candidate region before the correction.

The RCNN4SPTL adjusts the aspect ratio of region proposals for shape features of foreign objects on power transmission lines. So the RCNN4SPTL changes the aspect ratios of 1:1, 1:2, and 2:1 into: 1:1,2:1,3:1, because balloons hanging on transmission lines are mostly thin and long in images.

The RPN loss function combines the category scores of candidate boxes with the correction parameters. Equation (7) defines the loss function.

$$L(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}} \sum{}_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum{}_i p_i^* L_{reg}(t_i, t_i^*) \tag{7}$$

where i is the sequence number of a region proposal, p_i is the prediction confidence of a target in the i-th candidate area. $p_i^* = 1$ indicates that the i-th candidate region contains the object, and $p_i^* = 0$ indicates the i-th candidate box does not contain the object. t_i is the prediction correction parameter of the candidate, t_i^* is the correction parameter of the region proposal corresponding to the real region. N_{cls} and N_{reg} normalize the two sub-items in the formula (7) . is used to control the relative importance of two sub-items. L_{cls} () is the loss function of the prediction confidence as Equation 8:

$$L_{cls}(p_i, p_i^*) = -\log(p_i p_i^*) \tag{8}$$

L_{reg} () is the loss function of the modified parameter as (9).

$$L_{reg}(t_i, t_i^*) = \sum{}_{i\in\{x,y,w,h\}} smooth_{L1}(t_i - t_i^*) \tag{9}$$

Where $smooth_{L1}$ () is as equation (10).

$$smooth_{L_1}(x) = \{0\ .5x^2, |x| \le 1 \qquad |x| - 0.5, |x| > 1 \tag{10}$$

The formulas for calculating t_x^* , t_y^* , t_w^* , and t_h^* are as (11) , (12) , (13) , (14) respectively.

$$t_x^* = \frac{x^* - x_a}{w_a} \tag{11}$$

$$t_y^* = \frac{y^* - y_a}{h_a} \tag{12}$$

$$t_w^* = log\left(\frac{w^*}{w_a}\right) \tag{13}$$

$$t_h^* = log\left(\frac{h^*}{h_a}\right) \tag{14}$$

where x^* and y^* represent the horizontal and vertical coordinates of the center point of a real region, w^* and h^* are the width and height of the real region. x_a , y_a , w_a and h_a respectively represent the corresponding coordinates of a candidate region.

2.4. End-to-end joint training

The Faster RCNN uses alternating training. First the model is pre-trained on ImageNet to initialize shared convolutional network and then train the RPN. Next, the shared convolutional network is initialized using the pre-trained model on ImageNet, and train the classification regression network. Then the trained shared convolutional network and the classification regression network part parameters are fixed, and we start to train the RPN network. Finally the Faster RCNN initialize the entire network using the parameters trained in the previous step, the shared convolutional network and the RPNs parameters are kept unchanged and the Faster RCNN train the classification regression network.

We can see that the alternating training implies the feature sharing is actually a pseudo-sharing, which reduces the performance of the network. Therefore the RCNN4SPTL adopts the end-to-end joint training and treats the RPN and the classification regression network as a whole, and trains them simultaneously.

Firstly, the ImageNet pre-trained model is used to initialize the first two fully connected layers of the classification regression network and the shared convolutional neural network. The RCNN4SPTL randomly initializes the other layers using a Gaussian distribution with a mean of 0 and a standard deviation of 0.01 and perform end-to-ends Fine-tuning. In this training, the RPN and the classification regression network work together to train the shared convolutional neural network so that the RCNN4SPTL can simultaneously learn the required features. This kind of training can improve the performance to get a better model.

3. Image Preprocessing

The scale of training set affects model performance. The larger a training set, the better the deep learning model detects. Therefore we need to increase the size of training samples. The RCNN4SPTL adopts the following image preprocessing steps: image flipping, scaling and rotation to expand training sets. This study uses left and right flips; zooms all images to 400400 pixels. The RCNN4SPTL rotates images with 20 degrees, 100 degrees and 220 degree counterclockwise respectively to make the RCNN4SPTL invariant. Fig. 3 shows some examples of preprocessed images. Fig. 3 (a) shows the original images, and Fig. 3 (b) shows the pre-processed images, which are operated by image flipping flip, 20-degree rotation, and zooming.

(a)

(b)

Fig. 3. Preprocessed images

4. Evaluation

In order to evaluate the effectiveness of our approach, we use the following hardware for model training: NVIDIA GeForce GTX 1080TI with Intel i7@2.40GHz x 6 (6 cores) and 16GB RAM.

4.1. Data set

There are 5,000 training sample images in this experiment. Among them, there are 2000 films, 1000 balloons and 2000 kites. The test data set has 500 images, including 200 films, 100 balloons, and 200 kites. The example data set is shown in Fig. 4. The training set is manually labeled and processed. We fine-tune the hyper parameters of the

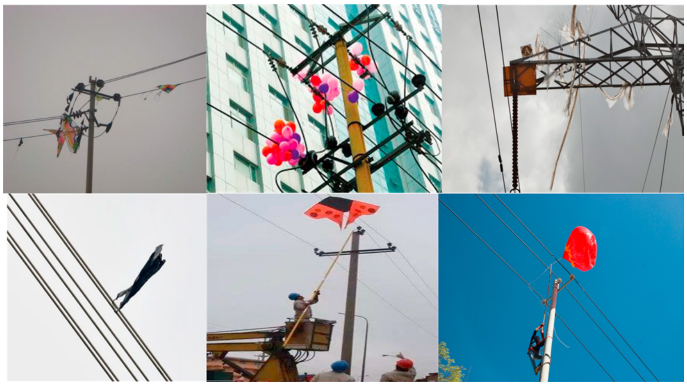

Fig. 4. Data set examples

RCNN4SPTL and then input the training set into the network for a limited number of iterative training. Finally, we utilize test sets to test the performance of the trained models, and we show the results in the following section.

4.2. Experimental results and analysis

Table 1. Performance comparison

Methods	Precision	Recall	Detection speed:Time per frame(*s*)
Hough Transform	0.3642	0.3918	5.52
Faster RCNN	0.7089	0.7275	0.30
Mask RCNN	0.7369	0.7514	0.28
FOTL RCNN	0.8601	0.8843	0.23

Table 1 shows the precision and recall rate of the test results. It demonstrates that the RCNN4SPTL has better detection performance on speed, precision and recall. The RCNN4SPTL is more suitable for detecting foreign objects

(a)

(b)

Fig. 5. Object detection results by RCNN4SPTL and Faster RCNN

than original Faster RCNN, in the case of detecting foreign objects on transmission lines. Figure 6 shows the results by RCNN4SPTL and Faster RCNN on detecting balloons, kites, and films. The test pictures are all from real scenes. Fig. 5 (a) lists the detection result using the RCNN4SPTL, and Fig. 5 (b) presents the result of detection using the Faster RCNN. It shows that the RCNN4SPTL recognizes foreign objects with a higher degree of confidence.

5. Conclusion

It is significant to detect and remove the foreign objects on power transmission lines in time. In this study, we first extend data sets using the specify image enhancement technology: image flipping, scaling and rotation. Then we propose the RCNN4SPTL network, which optimizes the shared convolutional network and the size ratio of the region proposals, according to the shape characteristics of the foreign objects of the power transmission line. Finally, we train the RCNN4SPTL using end-to-end joint training with 20000 iterations. The experimental results show that theRCNN4SPTL is more suitable for the accurate identification of foreign objects on transmission lines than the traditional Faster RCNN. TheRCNN4SPTL has faster detection speed, better recognition performance.

Acknowledgment

This research was supported in part by the National Major Science and Technology Project (2017ZX05013-002), the Key Research Program of Shandong Province (2017GGX10140) and in part by the Fundamental Research Funds for the Central Universities (2015020031), National Natural Science Foundation of China (61309024).

References

[1] Song Xuankun(2016) “Summary of China’s smart grid technology development practice ”,electric power construction,(7): 1-11.
[2] Wang J and Shao J(2017) “Image recognition of icing thickness on power transmission lines based on a least squares hough transform,”Energies,10(4): 415.
[3] Baker, L., Mills, S., Langlotz, T. and Rathbone, C(2016)“Power line detection using hough transform and line tracing techniques,” in Image and Vision Computing New Zealand (IVCNZ), 2016 International Conference on (pp. 1-6). IEEE.
[4] Zhao Yongsheng, Xu Haiqing and Wu Ligang(2017)“Application of hough line transform to foreign object recognition of transmission lines,”Digital Technology and Application,(3): 127-129.
[5] Zhou Feng and Renwei Xin(2018)“Image classification and feature extraction of transmission lines based on color space variables,”power system protection and control,46(5): 89-98.
[6] Huang Dongfang,(2016)“Research on traverse identification in transmission line images based on an improved Hough transform,”Guangxi University.
[7] Huang Xifei, Liu Baihong and Su Liangliang(2018)“A method for removing salt and pepper noise from images based on multi-window median filtering and iterative Gaussian filtering ,”(3): 77-78.
[8] Qiu Zhiqi(2015)“Image denoising based on median filter and wavelet transform,” North China University of science and technology.
[9] Yu Yonglong(2015)“Combined with bilateral filtering and dark channel image fogging algorithm and its application research,” Nanchang Aviation University.
[10] Ye, H., Yan, S. and Huang, P.(2017)“2D Otsu image segmentation based on cellular genetic algorithm,” in Communication Software and Networks (ICCSN), 2017 IEEE 9th International Conference on (pp. 1313-1316). IEEE.
[11] Cha Y J, You K and Choi W.,(2016)“Vision-based detection of loosened bolts using the Hough transform and support vector machines,” Automation in Construction,71: 181-188.
[12] Kong Yuanyuan.(2018)“Brain tumor segmentation method based on Hough transform localization and genetic algorithm,” Nanchang Aviation University.
[13] Ren X, Du S and Zheng Y(2017)“Parallel RCNN: A deep learning method for people detection using RGB-D images,” Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017 10th International Congress on. IEEE, 2017: 1-6.
[14] Li Y, Stevenson R L(2017)“Multimodal image registration with line segments by selective search,”IEEE transactions on cybernetics,47(5): 1285-1298..
[15] Purkait P, Zhao C and Zach C(2017)“SPP-Net: Deep absolute pose regression with synthetic views,” arXiv preprint arXiv:1712.03452, 2017.
[16] Qian, R., Liu, Q., Yue, Y., Coenen, F. and Zhang, B.(2016)“Road surface traffic sign detection with hybrid region proposal and fast R-CNN,” in Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016 12th International Conference on (pp. 555-559). IEEE.
[17] Eggert, C., Brehm, S., Winschel, A., Zecha, D. and Lienhart, R. (2017)“A closer look: Small object detection in faster R-CNN,” in Multimedia and Expo (ICME), 2017 IEEE International Conference on (pp. 421-426). IEEE.