Measurement 138 (2019) 379–385

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

Intelligent fault detection of high voltage line based on the Faster R-CNN

Xusheng Lei *, Zhehao Sui

School of Instrument Science and Opto-electronics Engineering, Beihang University, Beijing, China

ARTICLE INFO

Article history:
Received 7 June 2018
Received in revised form 27 October 2018
Accepted 26 January 2019
Available online 31 January 2019

Keywords:
Power line fault detection
Insulator
Bird nest
Convolution neural network
Faster R-CNN

ABSTRACT

To realize intelligent fault detection of high voltage line, a deep convolution neural network method based on Faster R-CNN method is proposed to locate the broken insulators and bird nests. With the region proposal network, the Faster R-CNN chooses a random region in the features of the image as the proposal region, and trains them to get the corresponding category and location for a certain component in the image. Since the internal and regional features of the image can be learned, the Faster R-CNN method transforms the problem of target classification into the problem of target detection and recognition. Based on the ResNet-101 network model, the damage of insulators and bird nests in the electric power line can be located effectively.

1. Introduction

Insulators are widely used for conductor insulation and mechanical support of electrical equipment in electric power systems [1]. Insulators can fix the busbar and live conductor, and provide enough distance among the conductors, which are important parts of the transmission line. However, insulators are very prone to be damaged under long term of rain, sun, climate change and chemical corrosion. The bird nests contain the wire, thatch, vines, cloth and excrement, which can lead to insulator flash over, short circuit and other failures. It may cause serious damage to the reliable operation of the grid. Therefore, insulators and bird nests are two typical faults for the high voltage line inspection.

With the quickly increment of high voltage transmission lines, the traditional manual inspection cannot finish the common inspection tasks. Based on the visible light camera loads, the images for the certain components on the high voltage line can be taken by the unmanned aerial vehicle (UAV). It can save cost and time effectively [2]. Therefore, using UAV as inspection tool to locate component faults has become a new trend for high voltage line inspection. Since the power lines are located in different environments, the background of the images taken by the UAV are very complexity. How to extract and locate the insulator failure and the bird's nest from the mass aerial images has become a bottleneck of the UAV power inspection.

The contour extraction, color features, texture features and machine learning methods are often applied to locate the insulators in the high voltage line in the past years. Markus used the circular structure feature to detect the insulator [3]. Based on the co-occurrence matrix of the edges, Khalay located the insulator successfully [4]. However, due to the complexity of the background of the images, the variability of the shooting angle leads to the uncertainty of the insulator contour. Based on the local features and spatial orders for aerial images, Liao proposed a robust insulator detection algorithm to realize high performance identification [5]. Focusing on green component in the HIS space, Lin used the maximum entropy threshold to locate the insulators [6]. But the method has certain limitation to achieve the ideal effect for the white, red, brown and black porcelain insulators. With grayscale co-occurrence matrix, Wu extracted insulator texture features from the aerial image [7]. Li used the MPEG -7 texture feature to detect the insulator [8]. But the pseudo targets with similar texture features may cause false location. Although the contour extraction, color features, texture features methods can develop the image recognition scheme, the hand-designed features are often time-consuming. It requires professional knowledge due to the diversity of features.

Machine learning and artificial neural network method can avoid the selection of hand-designed features. With the target suggestion and structure searching, Jabid used sliding window to estimate the insulator region [9,10]. But the algorithm is a time-consuming. Based on the neural network, Maraaba realized the classification of local insulator to evaluate the level of contamination [11]. However, it cannot locate the certain position of insulator in the image.

With the special network model structure, deep learning method learns the internal features of image rather than the single

* Corresponding author.
E-mail address: yushangtianxia@163.com (X. Lei).

https://doi.org/10.1016/j.measurement.2019.01.072
0263-2241/

feature of artificial design, which can enhance the image recognition and classification accuracy and stability. Since the common background and the color texture of aerial images are very complexity, a deep convolution neural network based on faster regions with convolution neural network features (R-CNN) method is proposed to locate the broken insulators and bird nests. It transforms the problem of target classification into the problem of target detection and recognition. Using the method of transfer learning, the small training samples can learn the basic features of the target, and locate the fault area in the image accurately according the learning results. Therefore, it can locate the broken insulator and bird nests effectively.

The rest of the paper is organized as follows. The structure of convolution neural network is introduced in Section II. Section III presents the power failure detection method based on Faster R-CNN. Section IV presents the experiments and results, followed by concluding remarks in Section V.

2. The basic structure of convolution neural network

As shown in Fig. 1, a typical convolutional neural network is mainly composed of the input layer, the convolution layer, the sampling layer and the full connection layer [12,13].

For the field of image recognition, the input layer of the convolution network is usually the original image, which is defined as X.

The convolution layer extracts the specific features of the image by the different sizes of convolution kernel. H_i is used to represent the feature map of the i-th layer of the convolutional neural network. The H_i can be generated as follows ($H_0 = X$):

$$H_i = f(H_{i-1} W_i + b_i) \tag{1}$$

where H_i is the feature map of the current network layer, H_{i-1} is the convolution feature of the previous layer, W_i is the weight of the i-th layer, and b_i is the offset vector of the i-th layer, f(x) is a nonlinear sigmoid function, and the symbol "$\otimes$" is the convolution operation.

The sampling layer, also known as the pooling layer, generally follows the convolution layer. The sampling layer reduces the dimension of the feature map and maintains the scale invariance of the feature map. Therefore, it can reduce the training parameters and calculation process.

The full connection layers play as the classifier in the whole network, which make a classification for the extracted features to obtain the corresponding probability of the outputs [14]. And the largest probability value is used as the prediction result of the neural network.

The training objective of the convolution neural network is to minimize the loss function of the network. The difference between the value of the result Y_i and the expected value $\widehat{Y}_i$ calculated by the input H_0 will be used to adjust the weight and residual error by the backpropagation algorithm.

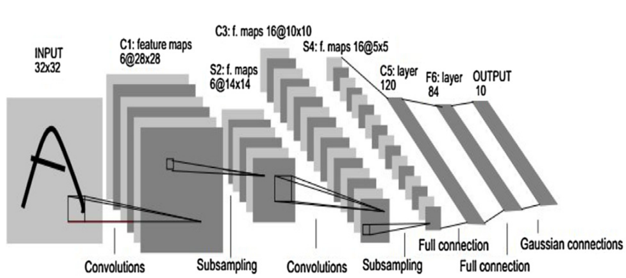

Fig. 1. A typical structure of a convolution neural network.

The general convolution neural network is composed of the convolution layer and the fully connected layer. The convolution layer is used to extract the features of the image, such as color, contour and types. And then the parameters of the convolution layer are adjusted by the backpropagation algorithm to learn the features of the image in the training process. The output of the convolution layer is the input of the fully connected layer, and the number of neurons in the fully connected layer corresponds to the classification of images. The convolution neural network structure can learn the features of the image, and make the corresponding classification for the whole image. But it cannot identify the category and location for a certain component in the images. However, fault location is to locate the failure component in the image rather than find the components in the images. Therefore, the general convolution neural network has limitation in the searching process.

3. Power failure detection based on the Faster R-CNN

The Faster R-CNN is a type of convolution neural network which is improved from R-CNN [15,16]. Based on the region proposal network, the Faster R-CNN chooses a random region of the image as the proposal region, and trains them to get the corresponding category and location for a certain component in the image. Therefore, it can realize high performance feature identification of the insulator failure and the bird nests from the mass aerial images. Compared with the traditional selective search method, the Faster R-CNN can break the bottleneck problem of huge cost in computation because the RPN can generate the corresponding proposal regions [17]. Therefore, it makes the real-time identification possible. Moreover, with the adaptive scale pooling layer, the Faster R-CNN can adapt to random image and tune the entire network to improve the accuracy of deep network identification.

The Faster R-CNN method not only breaks the time bottleneck of the calculation, but also ensures the ideal recognition rate. Therefore, the Faster R-CNN identification method is proposed to extract the features of the insulator and the nest to identify the target, the overall frame shown in Fig. 2.

The Faster R-CNN method is composed by two CNN networks. The Fast R-CNN detection network is in the upper half of the flow chart and the regional proposal network (RPN) is in the lower half of the flow chart [18]. The RPN samples the random region information of the image as the proposal regions, and train them to determine the areas that may contain the target [19]. The Fast R-CNN detection network further processes the area information collected by the RPN network, determines the target category in

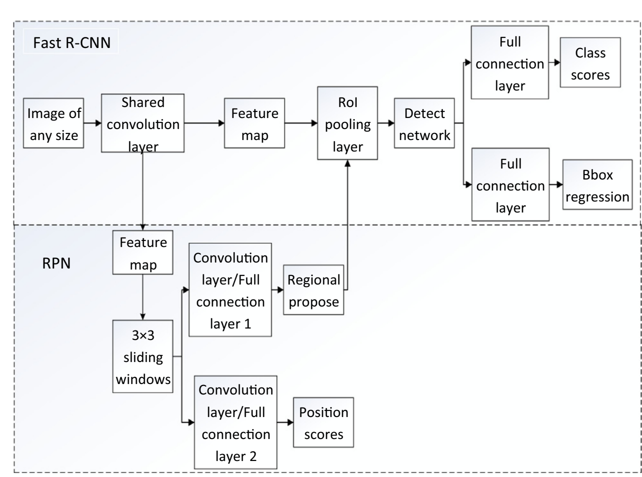

Fig. 2. The flowchart of the Faster R-CNN.

the area, and precisely adjusts the size of the area to locate the specific location of the target in the image.

The training process is shown in Fig. 3. Firstly, the parameters of the whole Faster R-CNN network were initialized with the pre-trained model, and then the RPN network was trained with our training data. The proposed region generated by trained RPN is then used to train the Fast R-CNN detection network. The RPN and the Fast R-CNN constitute a joint network, and the weight of the joint network is tuned by repeating the above process.

Based on the pre-training convolution neural network, the images including the insulators and nests are constructed as training data to optimize the parameters of the Faster R-CNN. The insulators and nests with certain categories and coordinates are marked in the images. With feature extraction network, the features of image are extracted. And then, some proposed regions can be generated by the RPN with the sliding window. Based on the Fast R-CNN, the category and coordinate of the planned insulators and the nests can be generated. With the comparison of the training data, the weight of the full connection layer of Fast R-CNN can be optimized.

3.1. The choice of feature selection network model

Since the selection of the convolution neural network for extracting features decides the final recognition accuracy rate in the training process, how to choose a suitable convolution neural network framework is very important in the training process.

In computer vision, the depth of the network is an important factor in achieving good results. With the increment of depth of the network, the features level increases corresponding. However, the large value of the depth will cause gradient vanishing phenomenon. If the layers behind the deep network are identity maps, the neural network model will degenerate into a shallow network, which can solve the problem generated by the increment of network layer. Therefore, the Residual Networks (ResNet) is proposed which is composed by a series of residual errors models [20], shown in Fig. 4.

The residual error function is defined as follows:

$$F(x) = H(x) - x \tag{2}$$

And then

$$\frac{\partial H(x)}{\partial x} = \frac{\partial F(x)}{\partial x} + 1 \tag{3}$$

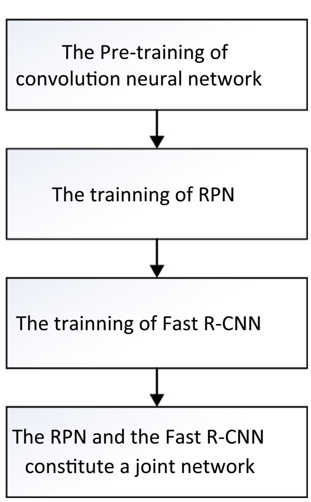

Fig. 3. The training process of Faster R-CNN.

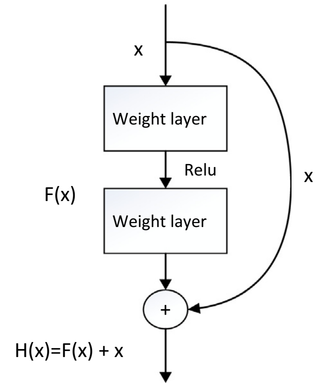

Fig. 4. The flowchart of residual structure.

In the process of gradient propagation, the gradient of the previous network is multiplied by the gradient of the latter network. In ResNet, its gradient is a number that is greater than 1. Therefore, it can avoid the gradient vanishing problem in the gradient propagation progress. The ResNet-101 with 101 layers network is chosen as the feature selection network model.

3.2. The pre-training CNN model

Due to the complexity of the high voltage line inspection, the real sampling images are not easy to obtained by the UAV system. But the huge number of samples are the basis of high precision training. Therefore, the transfer learning method is used to solve the problem [21]. It can use common data to get the pretrain model to construct the RPN network and the detection network.

The training data (about 100,000 images, 1000 classes) in the ImageNet is used to pre-train the Resnet-101 network model [22]. The first four convolutional neural networks in the ResNet-101 network are used as initialization parameters of the Faster R-CNN shared convolutional layer to extract features of image. The last level of the ResNet-101 network is used as the initialization parameter of the detection network in the Fast R-CNN.

3.3. The training of the RPN network

As the important part of the Faster R-CNN, RPN neural network generates different proposal regions and outputs them to the detection network to realize the identification for the proposed region. Its special network structure can promote the region extraction speed. At the same time, it can make the positioning more accurate based on the mark of the proposal region.

To locate the electric power failure in the inspection process, the images of the normal insulators, the damaged insulators and the bird nests are constructed the training set, which are marked by the coordinates of the minimum rectangular box containing the target and the category in the image.

With the convolution layer of the feature extraction network, the corresponding features of the images of the normal insulators, the damaged insulators and the bird nests in the training sets with random sizes can be generated. And then, a convolution layer based on the sliding operation is added to predict the position of the window containing the target. The convolution layer is composed by 9 benchmark widgets with three different sizes and three different proportions (1: 1, 1: 2, 2: 1). For each position of the feature map, the convolution operation is performed by opening a small window to obtain the multidimensional vector corresponding to the same position, which reflects the deep features in the small window of the position. And then, the probability value that the position of the small window belongs to the target/background and the position deviation of the small window relative to the real position window can be generated correspondingly.

When the information in the small window is recognized as the target, the proposal region is reserved and transformed to the subsequent network. If the information in the small window is recognized as the background, the proposal region is discarded. Since the small window traverses every position on the feature map, the electric power failure in the image can be covered by the sliding window because the position on the feature map corresponds to the position on the original image. Therefore, the insulators and the bird nests can be checked and transformed to follow-up testing network for the further judgments.

3.4. The training of the Fast R-CNN detection network

Based on the shared convolution layer in the ResNet network, the deep features of the images of the normal insulators, the damaged insulators and the bird nests are extracted by the feature extraction network. Combined with proposal region information of the RPN layer, the random size image is chosen as input for the pooling layer that is shown as ROI (Region of Interest). With the normalized process, the image with fixed size are transformed to the fifth stage of ResNet. And then, the output detection network is connected to a high dimensional feature vector by two full connection layers. The one full connection layer is used to realize classification for the target. Therefore, the probability that the candidate region box belongs to each category is scored. The other full connection layer is used to justify the position by the regional regression. With two translations and two zoom parameters, the candidate box can be adjusted to the real target.

In the training process, the parameters of neural network are adjusted by the loss function, which is defined as follows:

$$\mathrm{L}(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}}\sum_i L_{cls}(p_i, p_i^*) + \lambda\frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i, t_i^*) \tag{4}$$

where i is the index of an anchor in a mini-batch training and p_i is the predicted probability of anchor i. $t_i = \{t_x, t_y, t_w, t_h\}$ is a vector representing the 4 parameterized coordinates of the predicted bounding box. $p_i^*$is the ground-truth label. If the anchor is positive, the ground-truth label p_i^* is 1. If the anchor is negative, the p_i^* is 0. t_i^* is the ground-truth box associated with a positive anchor.

The classification loss $L_{cls}(p_i, p_i^*)$ is the log loss function for two classes, that is object or not object.

$$L_{cls}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \tag{5}$$

The regression loss function is

$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*) \tag{6}$$

where R is a robust loss function (smooth L1).

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & if|x| < 1 \\ |x| - 0.5 & otherwise \end{cases} \tag{7}$$

If the $p_i^* = 0$, the regression loss is disabled. If the $p_i^* = 1$, the regression loss is activated. The outputs of the cls and reg layers is composed by the $\{p_i\}$ and $\{t_i\}$ respectively. The two terms are normalized by N_{cls} and N_{reg} and weighted by a balancing parameterλ.

Based on the output of the loss function, the stochastic gradient descent method is used to adjust the parameters of the network to minimize the loss function in the backpropagation process. The flowchart is defined as follows:

1). Based on the Resnet-101 network model, the weight w and offset b of neural network are initialized.

2). The forward propagation calculation is performed according to the following three conditions.

2-1) If the current layer is the full connection layer:

$$a^{m,l} = \sigma(z^{m,l}) = \sigma\left(W^l a^{m,l-1} + b^l\right) \tag{8}$$

where $a^{m,l}$ represents the output value of the m image sample in the l layer, and $z^{m,l}$represents the value of the m image sample in the l layer neuron without activation.σ represents the activation function, and the Relu activation function is used in the network:

$$\mathrm{ReLU(x)} = \begin{cases} x(x > 0) \\ 0(x < 0) \end{cases} \tag{9}$$

2-2) If the current layer is the convolution layer:

$$a^{m,l} = \sigma(z^{m,l}) = \sigma\left(W^l \otimes a^{m,l-1} + b^l\right) \tag{10}$$

where $\otimes$ represents the convolution operation.

2-3) If the current layer is a pooling layer:

$$a^{m,l} = pool(a^{m,l-1}) \tag{11}$$

where the pool represents a reduced dimension operation.

3). For the output layer, layer L:

$$a^{m,l} = softmax(z^{m,l}) = softmax\left(W^l a^{m,l-1} + b^l\right) \tag{12}$$

$$\mathrm{softmax(z)}_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} (forj = 1, \cdots, K) \tag{13}$$

where K represents the dimension of z vector

4). Based on the loss function $J(\mathrm{W,b})$, the back propagation parameters $\delta^{m,l} = \frac{\partial J(\mathrm{W,b})}{\partial z^l}$ are defined according to the following three conditions

4-1) If the current layer is the full connection layer:

$$\delta^{m,l} = \left(w^{l+1}\right)^T \delta^{m,l+1} \times \sigma'(z^{m,l}) \tag{14}$$

where the $\times$represents the dot product.

4-2) If the current layer is the convolution layer:

$$\delta^{m,l} = \delta^{m,l+1} \otimes rot180(w^{l+1}) \times \sigma'(z^{m,l}) \tag{15}$$

where rot180 is the function of convolution kernel. The convolution kernel flips up and down, and then flips left and right.

4-3) If the current layer is a pooling layer:

$$\delta^{m,l} = unsample\left(\delta^{m,l+1}\right) \times \sigma'(z^{m,l}) \tag{16}$$

where unsample represents the anti-pooling operation.

5). Update the W^l and b^l of the l layer according to the following three conditions.

5-1) If the current layer is the full connection layer:

$$W^l = W^l - \alpha \sum_{m=1}^{n} \delta^{m,l} \left(a^{m,l-1}\right)^T \tag{17}$$

$$b^l = b^l - \alpha \sum_{m=1}^{n} \delta^{m,l} \tag{18}$$

where $\alpha$is the learning rate.

5-2) If the current layer is the convolution layer:

$$W^l = W^l - \alpha \sum_{m=1}^{n} \delta^{m,l} \otimes rot180\left(a^{m,l-1}\right)^T \tag{19}$$

$$b^l = b^l - \alpha \sum_{m=1}^{n} \sum_{u,v} \left(\delta^{m,l}\right)_{u,v} \tag{20}$$

5-3) If all the changes in W and b are less than the stop iteration threshold or the number of iterations required, then the cycle goes to step 6.

6). Output the linear relation matrix W and offset vector b of each hidden layers and output layers.

Using the marked insulators and the bird nests as references, the difference between the predicted region information can be generated by the detection network and the real target information. And then, the backpropagation algorithm is used to tune the weight and offset of the network. With the enough training, the Faster R-CNN can detect the accurate position and identification of the insulators and the bird nests.

4. Experimental results and analysis

To realize high performance fault location for the high voltage power line, the experiments are performed on a Windows PC with an Intel Core i7-7700 K, CPU (4.0 GHz), 16 GB DDR3 and an GTX1080Ti graphics card, 11 GB memory. The Caffe framework is chosen for the deep learning, which is a pure C ++/CUDA architecture [23].

The images taken by the UAV in the high voltage line inspection can be divided into three categories, including 120 images of broken insulators, 120 images of the bird nests, 320 images of the normal insulators. According to the ratio of 4: 1, the training set is composed by 448 images and the testing set is composed by 112 images. Due to the limitation of computation resource, the raw images are resized to 1000*750.

All images are marked by the minimum rectangular frame, saved by the xml format of VOC2007 [24]. It includes the four coordinates of the minimum rectangular and the category of the target in the box, shown in Fig. 5.

When the overlap area between recognized peripheral frame and the marked peripheral frame is more than 90% of the marked area, it is considered as a successful identification. The performance of the proposed method is evaluated by the values of precision and recall defined as follows:

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \tag{21}$$

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \tag{22}$$

where TP means true positive that is the number of successfully located targets.TP + FP is the total number of located objects, and TP + FN is the total number of actual targets.

The depth learning framework is based on the Caffe, and the ResNet-101 network model pre trained by the large data set ImageNet are chosen for the network model. For the ResNet-101 network, the residual network is divided into five large classes.

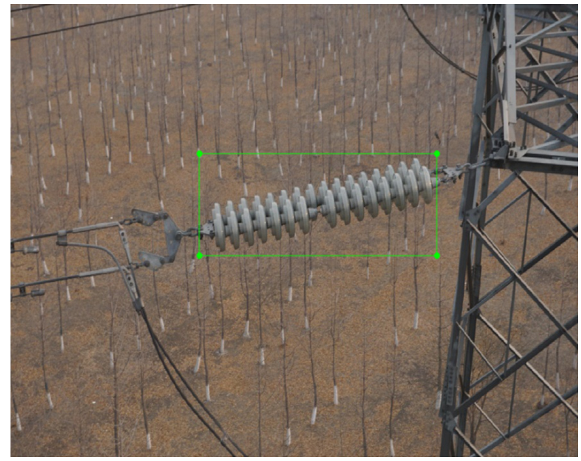

Fig. 5. The Labeled image.

The parameters of the first two classes were frozen to promote the training speed, and the rests of the classes were involved in the back propagation.

The training of Faster R-CNN is carried out using Momentum method [25]. We use a weight decay of 0.0001 and a momentum of 0.9. The learning rate is 0.001 for the first 40 k mini-batches and 0.0001 for the next 30 k.The number of mini-batch size is 8,

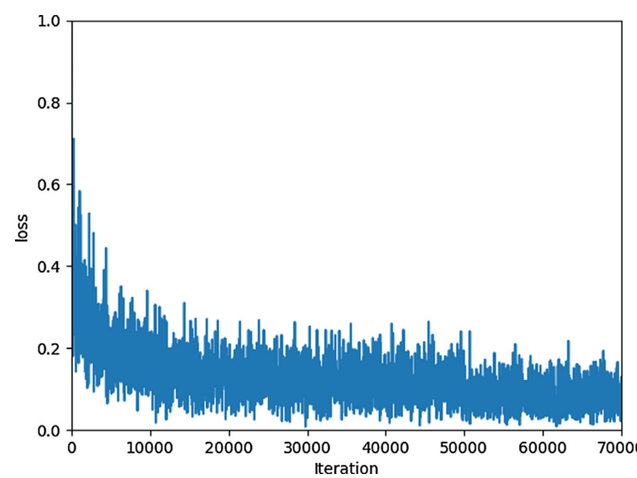

Fig. 6. The training loss of Faster R-CNN.

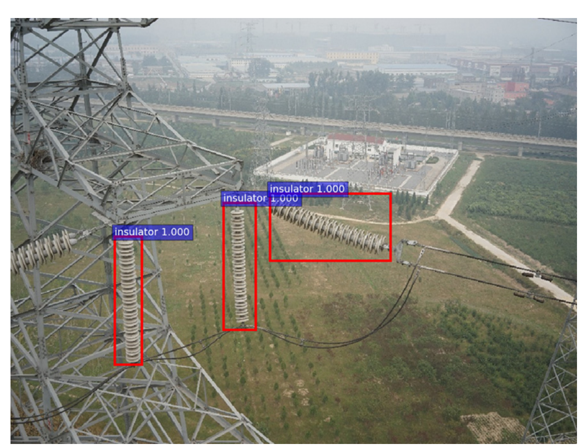

Fig. 7. The recognition results of normal insulators.

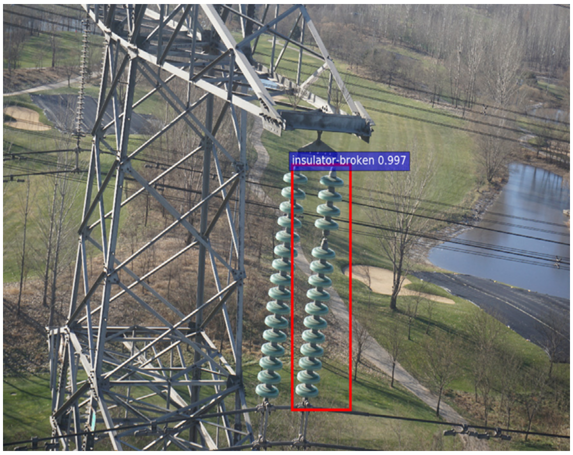

Fig. 8. The recognition results of damage insulator.

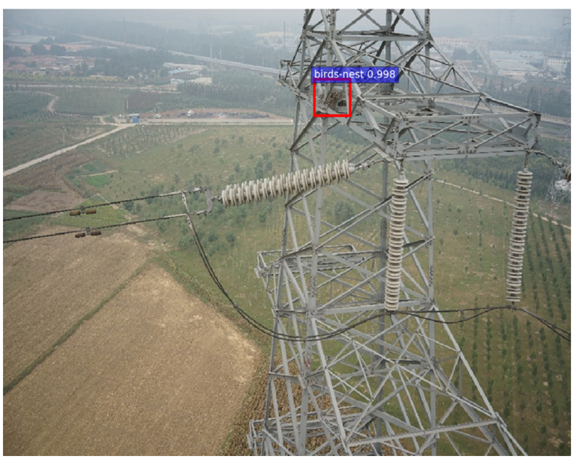

Fig. 9. The recognition results of Bird's nest.

Table 1
The Performance of different object detection models.

Model name	Pre-trained model	AP (average precision)	AR (average recall)	Speed (ms)
Faster R-CNN	resnet-101	0.976	0.979	201
	resnet-50	0.951	0.953	105
	inception-v2	0.948	0.941	96
	vgg-16	0.882	0.891	85
SSD	resnet-101	0.941	0.944	102
	resnet-50	0.936	0.912	50
	inception-v2	0.935	0.904	41
	vgg-16	0.861	0.891	18
YOLOv2	darknet-19	0.918	0.911	21
DPM		0.677	0.651	876

and the Intersection-over-Union(IoU) threshold for NMS is 0.7. The max step value is 70000. The training loss is plotted in Fig. 6.

Figs. 7–9 are the recognition images of the normal insulator, broken insulator, and bird nests images respectively. The detected targets are bounded by red boxes and the scores. The class name are shown on the above bottom of the boxes. With the Resnet-101 network, the value of average precision of the three classes is 0.976, and the average recall is 0.979. Besides, the average time cost on a single image is 201 ms. The proposed method can detect the power failures of high voltage line effectively.

To show the effectiveness of the proposed method, a series of comparison of different object detection methods have been done, including Faster R-CNN with different pre-trained CNN model, single shot multibox detector (SSD) with different pre-trained CNN models [26], YOLOv2 [27] and discriminatively trained part-based models (DPM) [28]. The comparison results are presented in Table 1.

For the four methods, DPM without the pre-trained model has a worst performance. The Faster R-CNN with resnet-101 model has the best identification performance, and the AP and AR are 0.976 and 0.979 respectively. The SSD with the vgg-16 model has the fastest speed, and the time is only 18 ms, but the AP and the AR are only 0.861 and 0.891 respectively.

To test the effectiveness of the Faster R-CNN with resnet-101 model, another 3000 images taken by the UAV in the high voltage line inspection have been used. With the automatic ordering program, the images are ordered and identified. The discrimination of broken insulators and bird nests surpass 97. 5%.

5. Conclusion

At present, most of the inspection image is analyzed by artificial classification. Due to the complexity of the background, artificial classification is prone to misjudgment. How to deal with the massive images has become the bottleneck of the quickly development of high voltage line.

To locate the insulator fault and nest automatically, a deep convolution neural network based on Faster R-CNN method is proposed to transform the problem of target classification into the problem of target detection and recognition. The Faster R-CNN with ResNet-101 pre-trained model has the best performance of accuracy can locate insulators and bird nests from massive images effectively. The average precision value is 0.976 and the average recall value is 0.979 in the experiment.

It is necessary to point that the time cost of single image is about 201 ms and the speed is not fast enough to detect video information in real time. Thus, we will improve our method to achieve the real-time detection with high accuracy in the future work.

Acknowledgments

This work was supported in part by the National Key Research and Development Plan (Grant number 2018YFB1107402), in part by the Beijing Science and Technology Plan (Grant number D171100006217003), in part by the National Science Foundation of China (Grant numbers 61873016 and 61633002), in part by the Fundamental Research Funds for the Central Universities (Grant number YWF-18-BJJ-214).

References

[1] U.K. Kalla, R. Suthar, K. Sharma, B. Signh, J. Ghotia, Power quality investigation in ceramic insulator, IEEE Trans. Ind. Appl. 54 (1) (2017) 121–134.
[2] F. Dong, X.S. Lei, W.S. Chou, A dynamic model and control method for a two-axis inertially stabilized platform, IEEE Trans. Ind. Electron. 64 (1) (2016) 432–439.
[3] M. Oberweger, A. Wendel, H. Bischof, Visual, recognition and fault detection for power line insulators, in: Proc. 19th Computer Vision Winter Workshop, 2014, pp. 1–8.
[4] L. Khalayli, H.A. Sagban, H. Shoman, K. Assaleh, A.E. Hag, Automatic inspection of outdoor insulators using image processing and intelligent techniques, in: Electrical Insulation Conference, 2013, pp. 206–209.
[5] S. Liao, J. An, A robust insulator detection algorithm based on local features and spatial orders for aerial images, IEEE Geosci. Remote Sens. Lett. 12 (5) (2017) 963–967.
[6] J. Lin, J. Han, F. Chen, X. Xu, Y. Wang, Defects detection of glass insulator based on color image, Power Syst. Technol. 35 (1) (2011) 127–133.

[7] Q. Wu, J. An, B. Lin, A texture segmentation algorithm based on PCA and global minimization active contour model for aerial insulator images, IEEE J. Selected Topics Appl. Earth Observ. Remote Sens. 5 (5) (2012) 1509–1518.
[8] W.G. Li, G.S. Ye, F. Huang, S.K. Wang, W.Z. Chang, Recognition of insulator based on developed MPEG-7 texture feature, High Voltage Apparatus 1 (2010) 265–268.
[9] C. Corinna, V. Vladimir, Support-vector networks, Mach. Learn. 20 (3) (1995) 273–297.
[10] T. Jabid, M.Z. Uddin, Rotation invariant power line insulator detection using local directional pattern and support vector machine, in: International Conference on Innovations in Science, Engineering and Technology, IEEE, 2017, pp. 1–4.
[11] L. Maraaba, Z. Alhamouz, H. Alduwaish, A neural network-based estimation of the level of contamination on high-voltage porcelain and glass insulators, Electric. Eng. 1 (2017) 1–10.
[12] Y. Lécun, L. Bottou, Y. Bengio, et al., Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[13] Y. Lu, S. Yi, N. Zeng, Y. Liu, Y. Zhang, Identification of rice diseases using deep convolutional neural networks, Neurocomputing 267 (2017) 378–384.
[14] A.A.M. Al-Saffar, H. Tao, M.A. Talab, Review of deep convolution neural network in image classification, in: International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications, IEEE, 2018, pp. 26–31.
[15] S.Q. Ren, K.M. He, R. Girshick, J. Sun, Faster R-CNN: towards real-time object detection with region proposal networks, in: International Conference on Neural Information Processing Systems, MIT Press, 2015, pp. 91–99.
[16] R. Girshick, J. Donahue, T. Darrell, et al., Rich feature hierarchies for accurate object detection and semantic segmentation, in: IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, 2014, pp. 580–587.
[17] J.R.R. Uijlings, K.E.A. van de Sande, T. Gevers, A.W.M. Smeulders, Selective search for object recognition, Int. J. Comput. Vision 104 (2) (2013) 154–171.
[18] R. Girshick, Fast R-CNN[C], in: IEEE International Conference on Computer Vision, 2015, pp. 1440–1448.
[19] S. Yu, Y. Wu, W. Li, W. Zeng, A model for fine-grained vehicle classification based on deep learning, Neurocomputing 257 (2017) 97–103.
[20] K.M. He, X.Y. Zhang, S.Q. Ren, J. Sun, Deep residual learning for image recognition, in: Computer Vision and Pattern Recognition, IEEE, 2016, pp. 770–778.
[21] K. Weiss, T.M. Khoshgoftaar, D.D. Wang, A survey of transfer learning, J. Big Data 3 (1) (2016) 9.
[22] O. Russakovsky, J. Deng, H. Su, et al., ImageNet large scale visual recognition challenge, Int. J. Comput. Vision 115 (3) (2015) 211–252.
[23] Y. Jia, E. Shelhamer, J. Donahue, et al., Caffe convolutional architecture for fast feature embedding, in: Acm International Conference on Multimedia, ACM, 2014, pp. 675–678.
[24] M. Everingham, L. Gool, C.K. Williams, A. Zisserman, The pascal visual object classes (VOC) challenge, Int. J. Comput. Vision 88 (2) (2010) 303–338.
[25] I. Sutskever, J. Martens, G. Dahl, et al., On the importance of initialization and momentum in deep learning, in: International Conference on International Conference on Machine Learning, JMLR.org, 2013, p. 1139.
[26] W. Liu, D. Anguelov, D. Erhan, et al., SSD: single shot multibox detector, in: Computer Vision – ECCV 2016, Springer International Publishing, 2016, pp. 21–37.
[27] J. Redmon, A. Farhadi, YOLO9000: better, faster, stronger, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 7263–7271.
[28] P.F. Felzenszwalb, R.B. Girshick, D. Mcallester, et al., Object detection with discriminatively trained part-based models, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1627–1645.