Defect diagnosis technology of typical components on transmission line based on Fully Convolutional Network

LI Zhenyu1,2,*, WANG Wanguo1,2, LI Tao1, LIU Guangxiu1,2, LI Zengwei3, TIAN Yuan3

1 State Grid Intelligent Technology Co., Ltd, Jinan, 250101, China.
2 State Grid Shandong Electric Power Company Electric Power Research Institute, Jinan, 250001, China.
3 State Grid Shandong Electric Power Company, Jinan, 250001, China.

* a723582833@126.com

Abstract. This paper presents an intelligent defect diagnosis algorithm based on fully convolutional neural network for typical transmission line components. Based on the region-based fully convolutional neural network algorithm and combined with deformable convolution, feature context fusion, clustering, the feature expression ability of neural network is improved. Through the improvement, the algorithm adapts to the object deformation and scale difference. In addition, the training effect is improved by improving the sample labeling strategy. In this paper, the defect diagnosis of four kinds of transmission line components (insulator, vibration damper, grading ring, wire clamp) is realized.

1. Introduction
Transmission line is an important part of power system. The safe and reliable operation of transmission line is of great significance to the stable development of the country. With the development of power system, the scale of transmission line is increasing. The workload of operation and maintenance of transmission line is also increasing. The traditional transmission line inspection mainly relies on the manual field observation. It is carried out by using the naked eye or the telescope. The efficiency is low and the inspection requirement is difficult to meet. In recent years, with the rapid development of unmanned aerial vehicle (UAV), image processing and other technologies, the use of UAV to inspect transmission lines has been widely used. It overcomes the disadvantages of low efficiency and environmental factors. The inspection unmanned aerial vehicles carry visible light cameras and other equipment to collect comprehensive information on transmission lines and acquire a large number of inspection images. At present, the inspection images collected by the UAV are mainly viewed in the manual mode to judge whether the transmission line is defective or not. Because of the huge amount of image data, it requires to consume a lot of manpower and material resources. Therefore, the intelligent transmission line defect diagnosis technology has important research significance.

Transmission line defects are characterized by various types, large scale differences and complex field environment, which greatly increases the difficulty of diagnosis. The intelligent transmission line defect diagnosis technology is inseparable from the expression of the image feature. The traditional feature expression needs to be designed manually. The expression ability is limited and the operation is complicated, such as SIFT, SURF, LOG, HOG and other operators. In recent years, with the
development of deep learning technology, through the design of neural network, features have the ability of self-learning. Especially the development of convolutional neural network has made great progress in object detection technology. The neural network obtains the effective expression characteristics through the self-learning of the sample and avoids the disadvantages of the traditional method.

There are many components in the transmission line, and the defects are in various forms. At present, the defect diagnosis method is more specific to a particular component, and the coverage is narrow. For example, in the paper[1], a method of insulator recognition and defect diagnosis based on shape perception is proposed, which can diagnose insulators according to the shape and structure characteristics of insulator strings. In the paper[2], a method of insulator location and self-explosion defect diagnosis based on classifiers and distance criteria is proposed. First, the insulator is positioned by using the adaboost classifier, and then the self-explosion defect is judged by using the Euclidean distance of the adjacent insulator. The paper[3] proposes an identification method for foreign object of transmission line based on ORB algorithm. The frame difference method is used to identify foreign objects and tracked with feature points, which is suitable for video scenes. In the application of deep learning technology, the paper[4] proposes an RCNN-based UAV inspection image component recognition method, but lacks diagnostic research. The papers [5,6] use multi-mode (visible light, infrared and ultraviolet) fusion method to collect different feature information, and use neural network for classification, which are quite complex. North China electric power university has published a series of papers[7,8] on insulator state identification, which use depth convolution neural network to extract multi-scale features of insulators and use support vector machine (SVM) to classify insulators.

![Figure 1. The Various forms of transmission line component defects](image)

This paper studies the defect diagnosis of typical transmission line components (insulator, vibration damper, grading ring, wire clamp), and proposes an algorithm of defect diagnosis based on fully convolutional network. By improving the sample labeling strategy and using the methods of deformable convolution, feature context fusion and clustering, the defect diagnosis of typical transmission line components is realized.

2. Region-based fully convolutional neural network algorithm

At present, the traditional CNN network is effective in object classification, but not in object detection. The typical Faster R-CNN [9] object detection algorithm is all convolution operation before ROI pooling layer, and has translation invariance. However, the network structure after ROI poaching has no translation invariance. Therefore, the fully convolutional neural network algorithm(R-FCN)[10] is used to integrate the object position information into ROI pooling layer.
R-FCN adopts popular object detection strategy, which also includes two steps: region proposal and region classification. The candidate regions are extracted by the region proposal network (RPN). Through the RPN network, get the region of interest. Then the region of interest is classified by R-FCN. R-FCN adds an additional convolution layer after the convolution layer shared with RPN, forming the form of total convolution network. So the input of R-FCN and RPN is the whole image.

The last convolution layer of R-FCN generates $k^2$ position sensitive score maps for each class on the whole image. There are $C$ classes of objects plus one background, so there is an output layer of $k^2(C + 1)$ channels.

Finally, R-FCN uses position-sensitive RoI pooling layer and selective pooling to give one score for per RoI. The input of ROI pooling operation is the stereo block corresponding to a RoI on the score map with size $k^2 \times (C + 1) \times W \times H$ ($W$ and $H$ are the width and height of ROI). And each color block only picks out one bin of the corresponding position. Combine $k \times k$ bin into a new solid block of size $(C + 1) \times W \times H$. For example, in the Fig3, the first yellow position-sensitive score maps only take the bin in the upper left corner, and the last light blue position-sensitive score maps only take the bin in the lower right corner. After all, all the bins are recombined into a thin solid block in the right half of the image. So the output of ROI pooling is a solid block of size $(C + 1) \times k \times k$.

R-FCN uses a voting mechanism to get scores for each class. $k^2 k$ bins are summed directly to obtain the score of each class which is done separately. The final score of each class is obtained by softmax and used to calculate the loss. The loss function of the R-FCN algorithm is similar to the faster RCNN, which is composed of the classification loss and the regression loss. Classification uses cross entropy loss functions and regression uses the $l1$-smooth loss function.

The loss function formula is shown below:

$$L(s, t_{x,y,w,h}) = L_{cls}(s_c) + \lambda[c^* > 0]L_{reg}(t, t^*)$$

$$L_{cls}(s_c) = -\log(s_c)$$

Where: $L_{cls}(s_c)$ is the cross entropy loss of classification. $L_{reg}(t, t^*)$ is the regression loss. $t = \{t_x, t_y, t_w, t_h\}$ is a vector representing the offset predicted by the network. $t^*$ is a vector that has the same dimension with $t$, indicating the offset of anchor box relative to ground truth box. $c^* > 0$ means that the regression loss is calculated only in the foreground, not in the background.

3. Algorithm improvement strategy

3.1. deformable convolution

Due to the difference of shooting angle of transmission line components, the same type of defect will have geometric deformation. The traditional convolution operation geometry is fixed. When convolution is carried out at a certain position of the input image. The fixed regular position is
sampled first, and then the image values sampled are convolved as the output of that position. It does not have the ability to adapt to geometric deformation and depends entirely on the diversity of the sample data itself. The algorithm is not robust to the defects of deformation.

In order to adapt to the geometric deformation of the objects, the deformable convolution operation is introduced into the convolution neural network. The deformable convolution used in this paper is to add an offset variable at the position of each sampling point. Through these variables, the convolution kernel can sample randomly near the current location, instead of being limited to the previous fixed rules. In this paper, the increased offsets are a part of the network structure, which are calculated by convolution units to realize the self-learning. After the offsets are introduced, the size and position of convolution kernel can be adjusted dynamically according to the image content that needs to be recognized at present. The intuitive effect is that the position of the convolution kernel sampling point at different positions will adaptively change according to the image content, so as to adapt to the geometric shape of different objects such as shape and size.

![Figure 3. Convolution form](image)

Figure 3 shows the sampling method of normal convolution and deformable convolution with the convolution kernel size 3x3. (a) represents the 9 points (green) sampled by the normal convolution law. (b) (c) (d) is the deformable convolution, and an offset (blue arrow) is added to the normal sampling coordinates.

3.2. Feature context fusion

Due to the different types of transmission line components, there are scale differences, so the transmission line defect diagnosis algorithm is required to adapt to the difference of defect objects. The traditional object detection network is to identify the objects at the last layer, and the small scale objects cannot be characterized due to the refinement of high-level features. The common solutions are to use image pyramid and multiple feature layers prediction, but each feature layer of these methods is isolated and the ability of feature expression is insufficient. Based on the structure of the fully convolutional neural network, this paper combines the FPN[11] feature context fusion method. At the same time of prediction on the multiple feature layers, the high level features and low level features are fused to enhance the expression ability of features for different scale objects.

The core structure of extracting feature context fusion information is shown in Fig5. Usually the three featureMaps on the left are called C1, C2, C3. The scales of C1, C2 and C3 are decreasing according to the two-fold relationship. For example, C1 is 128*128, C2 is 64*64, and C3 is 32*32. The three corresponding featureMaps on the right are called P1, P2, and P3. The P3 is obtained by a convolution kernel with a core number of 256 and a core size of 1 * 1. The P2 is the result of upsampling P3 and adding C2 to the convolution kernel of 1*1. The P1 is the same as P2. Then the P1, P2 and P3 are respectively processed by a 3*3 convolution kernel to eliminate the aliasing effect of up-sampling and generate the final feature maps.
3.3. Anchor size clustering

The regions of interest are obtained through the region proposal network (RPN), adopting the idea of anchor boxes. Anchor boxes are candidate boxes with adjustable size. As shown in the Fig6, three sizes and three proportions of anchor boxes are taken as examples. The three sizes are small (128 in blue), medium (256 in red) and large (512 in green), and the ratios are 1:1, 1:2 and 2:1 respectively. There are 9 combinations. When the algorithm is implemented, the base scale of anchor is usually set as base_size*scales, the default base_size=16, scales=(8,16,32). However, this default anchor boxes size and proportion method is difficult to meet the requirements of transmission line component defect diagnosis.

In this paper, the size and proportion of anchor boxes are adjusted by k-mean value clustering. The purpose of clustering is to make larger IOU value between anchor boxes and adjacent ground truth. Set \((x_i, y_i)\) as the center point of the ground truth box, and \((w_i, h_i)\) as the width and height of the ground truth box. The distance measurement formula is shown as follows:

\[
d(box, centroid) = 1 - IOU(box, centroid)
\]

(3)

The specific steps of clustering are as follows:

1) First of all, set k clustering center points \( (W_i, H_i), i \in \{1, 2, ..., k\} \). \((W_i, H_i)\) is the width and height of anchor boxes.

2) Calculate the distance \(d\) between each label box and each cluster center. In the calculation, the center point of each annotation box coincides with the clustering center, and then the label box is assigned to the nearest clustering center.

3) After all the label boxes have been assigned, the cluster center point is recalculated for each cluster. The formula is shown below:

\[
W_i' = \frac{1}{N_i} \sum w_i, \quad H_i' = \frac{1}{N_i} \sum h_i
\]

(4)

Where: \(N_i\) is the number of labeled boxes of the \(i\) cluster, which means the average of the width and height of all labeled boxes in the cluster.

Repeat steps 2 and 3 until the clustering center changes very little.

4) Finally, through cluster analysis, this paper set scales= (1, 2, 4, 8, 16, 32), ratio= (0.3, 0.5, 1, 2, 3).
3.4. Sample label strategy

For the defect diagnosis of transmission line components, one of the difficult aspects is that the same type of defects in the same component also has a variety of manifestations. If the same labeling method is adopted for different defect manifestations, the effective detection of the objects cannot be achieved. Taking the slip defect of vibration damper as an example, it may be that two dampers collide with each other or collision with spacer and other parts. It's also possible that a damper can slip out alone. In order to adapt to different forms of defects, according to different forms of defect, different labeling strategies are worked out in this paper. For the form of a collision between two dampers, mark the two dampers together, as shown in the red box of the Figure 6(a). For the form of a single slip of a damper, mark the damper with the wires on the upper and lower sides, as shown in the red box of the Figure 6(b). For the form of collision between damper and the spacer bar, label damper with the upper part of the spacer bar, as shown in the red box of the Figure 6(c). In this paper, three kinds of slip forms of damper are trained as different defect types to realize the diagnosis of various slip forms of damper.

![Figure 6. Different labeling forms of vibration damper](image)

4. Test results

In this paper, Resnet101 network model is adopted for training. Meanwhile, positive samples and negative samples of transmission line component defects are trained together. The training samples are divided into 11 categories, including: damper, damper damage, damper slip, damper bending, grading ring, abnormal grading ring, clamp, clamp tilt, insulator, insulator self-explosion and insulator pollution. More than 4300 training samples are collected, and 400 samples are selected for testing. When the threshold is set to 0.5, the AP values for the test samples are shown in the following Table 1.

| Recognized type                  | Label     | AP  |
|----------------------------------|-----------|-----|
| damper                           | fzfc      | 0.903|
| damper damage                    | fzfcsh    | 0.886|
| damper slip                      | fzfcy     | 0.841|
| damper bending                   | fzfcwq    | 0.715|
| grading ring                     | jyh       | 0.875|
| abnormal grading ring            | jyhyc     | 0.836|
| clamp                            | xj        | 0.856|
| clamp tilt                       | xjqx      | 0.750|
| insulator                        | jyz       | 0.883|
| insulator self-explosion          | jyz_zibao | 0.849|
| insulator pollution              | jyz_wuhui | 0.761|

From the test results, it can be seen that this paper has a better identification effect for the defects of damper damage, damper slip, abnormal grading ring and insulator self-explosion. All AP values are above 0.8, which has strong practical value. For the identification of normal components, the AP value is generally above 0.85 because of the relatively large number of samples. For the diagnosis of defects such as damper bending, clamp tilt and insulator pollution, the AP value is between 0.7 and 0.8. Further research is needed. The identification results of typical defects in this paper are shown in the Figure 7:
5. Conclusion
Aiming at the defect characteristics of transmission line, a fast multi-object detection model based on fully convolutional neural network is constructed. The sample annotation form is improved, and the deformable convolution method is adopted to adapt the deformation of the defect objects. In this paper, the feature context information fusion unit is adopted to optimize the network model structure. At the same time of prediction on the multiple feature layers, the feature information of higher layers is introduced, which is more robust to the difference of defect object scales. For types with a diagnostic AP of less than 0.8 in defective components, the main reason is that defect sample is insufficient. In the next step, the sample expansion method will be studied to improve the diagnosis effect of transmission line component.

References
[1] Zhang Jingjing, Han Jun, Zhao Yabo. Insulator recognition and defects detection based on shape perceptual[J]. Journal of Image and Graphics. 2014, 19(8).
[2] Shang Junping, Li Chuxin, Chen Liang. Location and detection for self-expelde insulator based on vision[J]. Journal of Electronic Measurement and Instrumentation. 2017, 31(6).
[3] JIAO Shengxi, WANG Haiyang. The Recognition and Tracking of Foreign Body on Transmission Line Based on ORB Algorithm[J]. Science Technology and Engineering. 2016, 16(27).
[4] WANG Wanguo, TIAN Bing, LIU Yue. Study on the Electrical Devices Detection in UAV Images based on Region Based Convolutional Neural Networks[J]. Journal of Geo-Information Science. 2017, 19(2).
[5] Jin Lijun, Jianyong Ai, Zhiren Tian, Kai Gao, and Hua Huang. Pollution state detection of insulators based on multisource imaging and information fusion. In Dielectrics (ICD), 2016 IEEE International Conference on, vol. 1, pp. 544-547. IEEE, 2016.
[6] Jin Lijun, and Da Zhang. Contamination grades recognition of ceramic insulators using fused features of infrared and ultraviolet images. Energies 8, no. 2 (2015): 837-858.
[7] ZHAI Yongjie, WANG Di, ZHAO Zhenbing. Recognition Method of Insulator Based on Object Proposals and Structure Research. Journal of North China Electric Power University. 2016, 43(4).

[8] ZHAO Zhenbing CUI Yaping. Research progress of visual detection methods for transmission line key components based on deep learning. Electric Power Science and Engineering. 2018, 34(3).

[9] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 2015:91-99.

[10] Dai K J, Y L. Object detection via region-based fully convolutional networks. arXiv preprint arXiv:1605.06409, 2016.

[11] Lin T Y, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection. IEEE Conference on Computer Vision & Pattern Recognition. 2017.