# Research on Transmission Line Defect Detection Based on Deep Learning

**Liu Wuneng, Liu Lilong, Luo Changbing, Li Bin**
*The Diqing Power Supply Bureau, Yunnan Power Grid Co., Ltd, Qujing, Yunnan, 650051, China*

***Abstract***

*At present, the inspection method of transmission lines in China is manual inspection assisted by unmanned aerial vehicles, which has a low degree of intelligence. In order to improve the intelligent level of patrol inspection, the main defects of transmission lines, such as insulator self-explosion and bird's nest, are taken as detection objects, aiming to explore a detection method of transmission line defects with high detection accuracy and high speed. Aiming at the problems of heavy workload and low efficiency of manual defect image recognition, deep learning technology is introduced into the defect recognition module of transmission line engineering acceptance. By integrating the optimized Faster-R CNN image recognition algorithm to learn and recognize the collected images, a lightweight transmission line defect detection method is constructed by combining depth separable convolution and SVD decomposition. Experimental results show that the effectiveness and reliability of the deep learning method in the identification and defect detection of high-voltage transmission line components are very high, and Faster-R CNN can achieve the identification speed of nearly 0.147 s per piece.*

***Keywords****: Deep learning, Faster-r cnn, Transmission lines, Defect detection*

## Ⅰ. Introduction

Safe operation of power grid is a major event related to people's livelihood. Regular inspection of power transmission lines is a heavy task. In recent years, aerial survey has gradually replaced the traditional manual inspection and robot inspection, and has been widely used in transmission line inspection [1-2]. In reference [3], according to the difference characteristics of aerial insulator images, a method of measuring the difference degree based on membership function is designed to judge whether the insulator state is normal or not. Literature [4] uses Pulse Coupled Neural Network (PCNN) to extract aerial insulator images and uses generalized Hough transform to identify and locate targets. In reference [5], by studying the principle of generalized Hough transform, an identification and positioning method for insulator contour is proposed.

At present, the deep learning theory shines brilliantly in the direction of machine vision, and it is feasible to use the deep neural network theory to realize the intelligent inspection technology of UAV. Smarter patrol inspection will replace traditional patrol inspection in a large area, greatly reducing labor cost and improving detection efficiency. Because of the particularity of patrol inspection task, it is necessary to research and design a neural network with fast detection speed and high detection accuracy to detect the main defects of transmission lines. In this paper, the VGG Net[10] network is reconstructed by using deep separable convolution, and SVD is used to decompose the final classified and regressed network coefficient matrix of Faster-R CNN [6-7]. Through the fusion of two steps, the network parameters and computation are reduced to realize the compression of the network model. The experimental results on the data set of transmission line defect detection show that the model size of the proposed method is significantly reduced, the detection computation is significantly reduced, and the defects in transmission lines can be effectively detected.

## Ⅱ. Faster-Rcnn Target Detection Network

Faster-RCNN network is the most used neural network in two-stage algorithm, which has higher accuracy and faster speed than similar networks. The detection process of Faster-R CNN is shown in Figure 1.

ISSN: 0010-8189

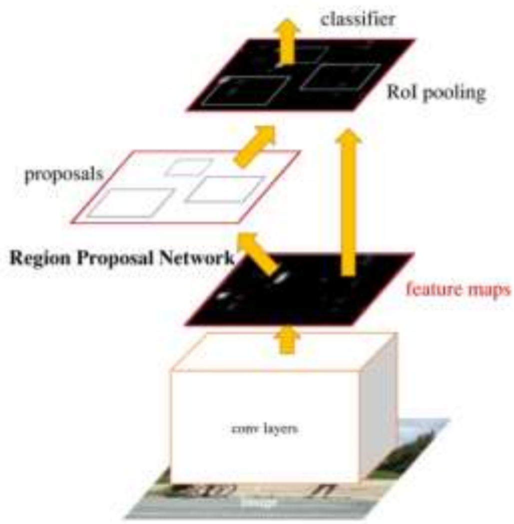


*Fig.1 Schematic Diagram of Detection Process of Faster-r Cnn Target Recognition Network*

The innovation of this network is to introduce the Region Proposal Network (RPN) network to locate the target preliminarily. The RPN network divides the feature map generated by convolution layer into different regions by using sliding windows, and uses anchor boxes with different horizontal and vertical ratios and areas in each divided region. Each anchor box is used as a candidate region, and the region is classified and judged. Then regress and fine-tune the coordinates of each anchor box in the region, and finally send the information tensor containing the target coordinates and scores to the Region of interest pooling (RIO) layer. Finally, the network locates and predicts the generated features carefully, and outputs the results of target detection. Compared with Faster-R CNN and other networks, this network accelerates the processing speed and prediction accuracy of the network.

## Ⅲ. Image Detection of Transmission Line Defects

### *A. Svd-Based Network Full Connection Layer Compression*

SVD is a model compression method that approximately represents the original matrix by extracting key features (corresponding to large singular values) [8]. The specific implementation method is to perform singular value decomposition on a matrix with larger dimension, and to approximate the original matrix by multiplying several matrices with smaller dimensions representing key features in the original matrix.

Let the dimension of matrix $W$ be n×m, and carry out SVD decomposition on $W$:

$$W = U\sum V^T \quad (1)$$

In which $U$ and $V$ are column orthogonal matrices of n×r and m×r respectively; $\sum$ is a diagonal matrix of singular values of r×r (non-negative singular values are arranged in descending order); $r\ \left(r \le \min\left(m, n\right)\right)$ is the rank of matrix $W$.

$$\Sigma = diag\left(\sigma_1, \sigma_2, \cdots \sigma_r\right) \quad (2)$$

In which: $diag\left(*\right)$ is a diagonal matrix; $\sigma_i$ is the $i$ th singular value.

ISSN: 0010-8189

The larger singular value corresponds to the more critical information component. The largest $k\left(k \le r\right)$ singular values of matrix $W$ are extracted and recorded as $\sum^{'}$. Let $U_k$ and $V_k$ be column submatrixes corresponding to the maximum $k$ singular values corresponding to $U$ and $V$, respectively, where,

$$\sum^{'} = diag\left(\sigma_1, \sigma_2, \cdots \sigma_k\right) \quad (3)$$

$$U^{'} = U_k \sum^{'\frac{1}{2}} \quad (4)$$

$$V^{'} = V_k \sum^{'\frac{1}{2}} \quad (5)$$

The matrix $W$ can be approximately expressed as:

$$W \approx U^{'}V^{'T} \quad (6)$$

With SVD decomposition approximation, the dimension of $W$ is reduced from n×m to $\left(n + m\right)k$.

In the convolutional neural network, it is found that the redundancy of the parameter matrix of the fully connected layer network is very large, that is, the rank of the parameter matrix of the fully connected layer network is very small. Convolution operation of convolutional neural network mainly affects the computation, while the full connection layer affects the size of network model storage. Therefore, SVD is used to decompose and compress the final all-connected layer network parameter matrix of defect detection model. The SVD decomposition of two-dimensional tensor is extended to three-dimensional convolution, and the initial matrix is expressed as the product of low-dimensional matrices by SVD decomposition.

***B. Convolution Neural Network Model***

Convolutional neural network model belongs to multi-layer neural network, each layer of neural network is composed of multiple two-dimensional planes, and each plane contains multiple neurons. Generally, it includes convolution operation and pooling operation [9]. The input of convolution neural network is digital image, and then convolution layer and pooling layer alternate. Convolutional neural network can reduce the number of parameters in two ways, namely local visual field perception and parameter sharing.

The feature map obtained by convolution layer is the result obtained by adding an offset to the input image and convolution kernel operation and then activating the function. Pooling layer can locally average the features in convolution layer. Generally, the size of pooling layer is n*n, and the maximum or average value of n*n adjacent cells in convolution layer corresponds to max-pooling and mean-pooling methods, which can reduce the resolution of feature map and reduce the sensitivity to displacement deformation. Figure 2 shows the principle of pooling operation.

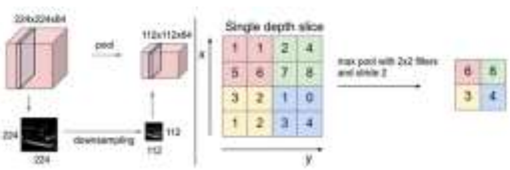


*Fig.2 Schematic Diagram of Pooling Operation*

ISSN: 0010-8189

www.converter-magazine.info

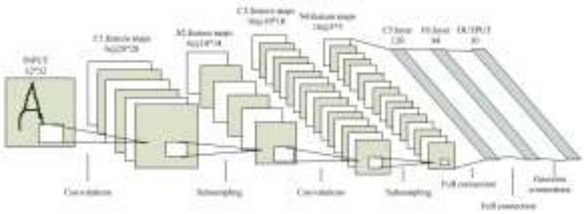

*Fig.3 The Network Structure Diagram of Lenet-5*

Deep convolutional neural networks are generally composed of convolution layer, pooling layer, full link layer and softmax layer. Fig. 3 is a network structure diagram of LENET-5, which mainly includes two groups of convolution and pooling layers. Then, the feature vector of the image is mapped to the classification output vector through the full link layer twice.

***C. Improvement of Network Structure***

In order to speed up the speed, depth separable convolution is used instead of ordinary convolution in multi-scale feature fusion pyramid structure. In this chapter, the modules and functions of skeleton network MobeliNetV2 are introduced in detail, and then the structure of the improved YOLOv3 network is introduced. the calculation amount of depth separable convolution is compared with ordinary convolution, and the ordinary convolution in feature fusion pyramid is replaced by depth separable convolution.

(1)MobeliNetV2 skeleton network

MobeliNetV2 network is a lightweight network released by Google Inc. The innovation of this network is linear bottleneck and inverted residual module. Using deep separable convolution can greatly reduce the network computation and speed up the network operation. The inverted residual module of the network has good feature extraction performance, which improves the detection performance of the network.

Generally speaking, with the increase of the depth of neural network, the features that can be extracted by the network will be more abstract and the performance will be better. However, experiments show that when the network depth reaches a certain number of layers, the detection accuracy of the network will decrease. This is because the training of deep neural network uses gradient back propagation algorithm. With the increase of network layers, the gradient of back propagation becomes smaller and smaller, which leads to the problem of gradient dispersion or gradient disappearance, and the front neural layer can not learn. We call this phenomenon network degradation. While the convolution operation in MobeliNetV2 network is depth-separable convolution, which reduces the computation, but the dimension of its output features is much smaller than ordinary convolution operation [10]. As shown in fig. 4, the dimension of the input feature map is smaller than that of ordinary convolution. if the residual network is calculated according to the method of first compressing and then expanding, the features extracted by the network will be very few. Therefore, we do the opposite. First, we enlarge the input feature channel by 6 times, that is, we upgrade the dimension. Finally, we use the linear bottleneck layer mentioned above to reduce the dimension. This method is called inverted residual module.

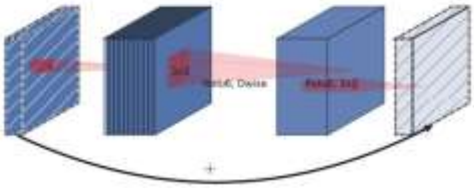

*Fig.4 Comparison Diagram of Feature Dimension Change between Residual Module and Inverted Residual Module*

Inverted residual module firstly uses 1*1 convolution and ReLU6 nonlinear active layer to enhance the dimension

ISSN: 0010-8189

of feature map, then uses 3*3 separable convolution and ReLU6 nonlinear active layer to extract features, finally uses 1*1 convolution linear bottleneck layer to compress features, and then adds the input and new output features for output.

(2) Improved YOLOv3 network structure

The structure of neural network determines the detection accuracy and speed of the network. In order to meet the requirements of high detection accuracy and fast running speed for the main defect detection tasks of transmission lines, the original YOLOv3 detection network was improved. In this paper, MobeliNetV2 network is used as the skeleton network of the network, and separable convolution is used to replace the ordinary convolution of the original network in the multi-scale feature pyramid to speed up the running speed of the network. The architecture diagram of the improved YOLOv3 network is shown in fig. 5.

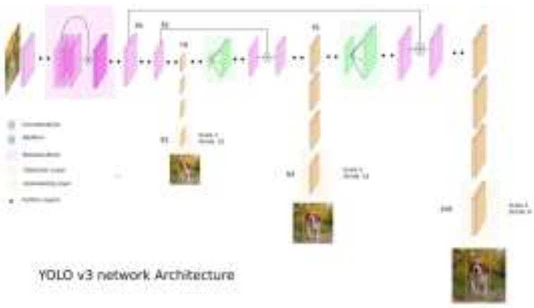


*Fig.5 Improved Yolov3 Target Detection Network Structure Diagram*

YOLOv3's Bounding Box is called Bounding Box, which draws lessons from the anchor mode of RPN network in Faster-R CNN, but does not want to manually design the horizontal and vertical ratio and size of the prior box according to prior knowledge, so bounding box proposes to predict four offsets of the bounding box for accurate target detection.

It is assumed that the coordinate of the upper left corner of a divided cell in the feature graph is $(C_x, C_y)$ , and the width and height of the Bounding Box are $P_w$ and $P_h$. Although the value of the Bounding Box is set in advance, because the bounding box adopts the offset mechanism, the finally produced coordinates and length and width values are learned, and the preset value has no influence on the positioning of the network. The coordinates and dimensions of the Bounding Box are shown in Equations (7) to (10).

$$b_x = \sigma(t_x) + c_x \quad (7)$$

$$b_y = \sigma(t_y) + c_y \quad (8)$$

$$b_w = p_w * e^{t_w} \quad (9)$$

$$b_h = p_h * e^{t_h} \quad (10)$$

In the formula, $b_x, b_y, b_w, b_h$ is the coordinate and width of Bounding Box predicted, $c_x, c_y, p_w, p_h$ is the set coordinate and width, $\sigma(t_x)$ and $\sigma(t_y)$ are the coordinate offset, $e^{t_w}$ and $e^{t_h}$ are the length-width scaling ratio.

ISSN: 0010-8189

Each feature map in YOLOv3 is divided into several small blocks (i.e., cell), each block predicts three Bounding Box, and each box contains three pieces of information, namely, the coordinates and width and height of each box, which are the four values introduced above, the target prediction score and the category prediction.

## Ⅳ. Data Set and Model Training

The data set in this paper mainly comes from the pictures taken by unmanned aerial vehicles during patrol inspection. The pictures cover four seasons: spring, summer, autumn and winter. The shooting locations are diverse, and the pictures collected by unmanned aerial vehicles have high resolution. In this paper, aiming at the problem of identification and defect diagnosis of transmission line components, five different types of components, namely, equalizing ring 1, equalizing ring 2, intact anti-vibration hammer, broken anti-vibration hammer and bird's nest, are selected as model training samples. Among them, there are 1000 training samples in each category, totaling 5000 training samples, and the size of each sample picture is 6500× 4500.

In this paper, based on mxnet framework, faster-rcnn is used to train the network model. ZF network, vgg16 network and resnet-101 network are used to initialize the pre-trained imagenet network in turn, and three models with different network layers are obtained. For each model training, the training times are 25 times, the batch size is 138, the learning rate is 0.001, the weight attenuation rate is 0.0005, and the number of candidate regions before and after non-maximum suppression is 7000 and 400 respectively. In this experiment, the accuracy rate, recall rate and missing recognition rate are used as the evaluation criteria of the model, in which the accuracy rate represents the ratio of the number of correctly recognized targets to all recognized targets; The recall rate represents the ratio of the number of correctly identified targets to the number of targets of all samples; The missing recognition rate indicates the ratio of the number of unrecognized targets to the number of targets in all samples.

## Ⅴ. Experimental Results and Analysis

### *A. Testing and Verification*

1 000 aerial images of unmanned aerial vehicles were used in the test of digital image defect recognition technology for transmission towers, including images of 20 kinds of objects, each of which has 25 defective and non-defective images. The test is divided into two parts: 1) Identify the equipment that needs defect diagnosis; 2) It is necessary to judge the defect type and precise location for each kind of equipment. Fig. 6 is the scene diagram of UAV assisted patrol.

*Fig.6 Unmanned Aerial Vehicle Assisted Inspection Site Map*

Tests show that for the first part, the multi-target object recognition algorithm based on Faster-Rcnn can recognize multiple objects from the same picture at the same time. And has the advantages of high speed and high recognition accuracy. It takes only 0.3~0.6s to process a high-definition picture with 24 million pixels on NvidiaTitanX. 20 kinds of objects are classified on ImageNet, and the accuracy is about 85%. Under the current experimental conditions, the recognition accuracy of common transmission equipment can reach 85.5%.

### *B. Influence of Convolution Kernel Size on Recognition Effect*

In view of the unsatisfactory recognition effect of shock hammer and the optimization of single picture recognition time, the optimal model Resnet-101 is selected for the experiment. The experimental direction is to modify the convolution kernel of its convolution structure, mainly aiming at the size of the first layer convolution kernel. According to the content in Section 1, it can be known that different convolution kernel sizes will affect the recognition accuracy and recognition time.

The convolution kernel size in the first convolution layer of ResNet101 model network is 7. According to the experimental idea, the convolution kernel size is gradually reduced for model training. Then the trained model is used to detect the samples, and the recall rate and recognition time are used as experimental indexes The number of samples is 500, and the experimental results are shown in Table 1.

*Table 1 Recall Rate Of Different Convolution Kernel Sizes*

| Convolution | Equalizing ring 1 | Equalizing ring 2 | Bird's nest | Good shock hammer | A broken shock hammer |
|---|---|---|---|---|---|
| 9*9 | 0.936 | 0.933 | 0.916 | 0.886 | 0.847 |
| 7*7 | 0.901 | 0.921 | 0.903 | 0.843 | 0.826 |
| 5*5 | 0.886 | 0.871 | 0.879 | 0.830 | 0.814 |
| 3*3 | 0.863 | 0.860 | 0.856 | 0.804 | 0.801 |

According to the experimental results, it can be seen that different convolution kernel sizes have an impact on the detection accuracy, and it can be obtained that with the decrease of convolution kernel size, the recall rate decreases continuously, because the large convolution kernel has a large receptive field, and the recognition accuracy of the convolution kernel with a large receptive field is also high.

### *C. Comparison of Different Algorithms*

On the COCO data set, besides the experiments of the original Faster-R CNN algorithm and lightweight compression algorithm, a comparative experiment with the lightweight network MobileNetV1 is also added. The feature extraction network of the original Faster-R CNN is replaced by MobileNetV1. The quantitative pairs of detection performance of the three algorithms are shown in Table 2.

*Table 2 Comparison of Detection Results of Different Algorithms on Coco Data Set*

| Algorithm | m AP/% | Parameter quantity/Million | Model size/MB | Testing time/s |
|---|---|---|---|---|
| Faster-R CNN | 25.31 | 138.66 | 1354.78 | 0.147 |
| MobileNetV1 | 19.80 | 6.71 | 66.89 | 0.036 |
| Lightweight compression algorithm | 21.04 | 6.02 | 54.21 | 0.028 |

It can be seen from table 2 that the computational complexity and model size of the network trained by the lightweight compression algorithm are reduced by two orders of magnitude compared with the original Faster-R CNN, and compared with the Faster-CNN algorithm fused with MobileNetV1, the lightweight compression algorithm is superior to MobileNetV1 in both detection accuracy and model size. This also shows the advantages of deep separable convolution and SVD fusion compression.

## Ⅵ. Conclusion

In this paper, focusing on the intellectualization of defect identification in the application of transmission line project acceptance, aiming at the specific scene of project acceptance, the typical defects in the project are studied and analyzed, and a calculation method based on Faster-R CNN is proposed. A lightweight compression algorithm is constructed by using the fusion compression of deep separable convolution and SVD decomposition. Compared with the original Faster-R CNN method, the lightweight compression algorithm reduces the network parameters and model size by two orders of magnitude without reducing the detection accuracy of COCO data sets; The experimental results in this paper show that deep learning can be well applied in component identification and defect detection of transmission lines, which can lay a good foundation for intelligent processing of UAV inspection images in the later period.

### Ⅶ. Acknowledgments

Based on the depth of visual transmission line defect feature acquisition and recognition research. No.: 051500kk52180011

### References

[1] Zhao B, Dai M, Li P, et al. Defect Detection Method for Electric Multiple Units Key Components Based on Deep Learning[J]. IEEE Access, no. 99, pp. 1-1, 2020.
[2] Li Zhaojian, Jiang Xiujuan, Zhu Zhengtao, et al. Bullet Defect Detection Method Based on Deep Learning[J]. combined machine tool and automatic machining technology, vol. 000, no. 009, pp. 102-106,110, 2019.
[3] He Zhiming, Peng Yanan. Progress in fabric defect detection based on deep learning[J]. Wool spinning technology, vol. 047, no. 008, pp. 83-88, 2019.
[4] Ma X, Kittikunakorn N, Sorman B, et al. Deep Learning Convolutional Neural Networks for Pharmaceutical Tablet Defect Detection[J]. Microscopy and Microanalysis, vol. 26, no. S2, pp. 1-5, 2020.
[5] Ma X, Kittikunakorn N, Sorman B, et al. Application of Deep Learning Convolutional Neural Networks for Internal Tablet Defect Detection: High Accuracy, Throughput, and Adaptability - ScienceDirect[J]. Journal of Pharmaceutical Sciences, vol. 109, no. 4, pp. 1547-1557, 2020.
[6] Tao Xian, Wei Hou, Xu De. Overview of surface defect detection methods based on deep learning. acta automatica sinica, vol. 47, no. 5, pp. 1017-1034, 2021.
[7] Deif S, Daneshmand M. Long Array of Microwave Sensors for Real-Time Coating Defect Detection[J]. IEEE Transactions on Microwave Theory and Techniques, no. 99, pp. 1-11, 2020.
[8] Khan A Q, Ullah Q, Sarwar M, et al. Transmission Line Fault Detection and Identification in an Interconnected Power Network using Phasor Measurement Units[J]. IFAC-PapersOnLine, vol. 51, no. 24, pp. 1356-1363, 2018.
[9] Xu Heng, Peng Shurong, Mao Yazhen, et al. Transmission Line Ice Thickness Detection Based on Image Processing[J]. Shaanxi Electric Power, vol. 045, no. 005, pp. 32-35, 201.
[10] Davis, Mark. Characterization of AlSi and AlSi-TiSi2 Metal-Semiconductor Contacts[J]. Journal of the Microelectronic Engineering Conference, vol. 24, no. 1, pp. 13-13, 2018.