Infrared image detection of insulators based on Centernet model

Xiaogang Li¹, Yongheng Ku², Yaohui Cui³*, Zhansheng Tian¹ and Wukui Sun¹

¹ Henan Jiuyu EPRI Electric Power Technology Co., Ltd, Zhengzhou, Henan, 450000, China
² State Grid Henan Electric Power Company, Zhengzhou, Henan, 450000, China
³ School of Electrical Engineering, Guangxi University, Nanning, Guangxi, 530004, China

*Corresponding author’s e-mail: lpcyh123@163.com

Abstract. The infrared image can reflect the temperature information inside the insulator, which is helpful to realize the follow-up fault diagnosis. However, insulators are usually located in remote areas, and the complex background will affect the accuracy of the algorithm model. Therefore, we established an infrared dataset of insulators. We introduced the Centernet network to infrared detection of insulators, and conducted training and testing on the data set. The results show that our proposed model has a significant improvement of 4% to 5% compared with several mainstream detection models. This can meet the needs of insulator fault detection.

1. Introduction

Insulators are important components to maintain the stability of the power system[1-3]. Since the substations and powerlines are mostly located in remote areas, the insulators are exposed to the natural environment for a long time, so they are susceptible to the influence of the harsh environment and cause failures, affecting the normal operation of the power system.

The traditional manual detection method consumes manpower and material resources, and the area where the powerline located is complicated, so the manual detection efficiency is not high. Since the birth of deep learning, people have continued to apply deep learning to various fields. Compared with manual detection and traditional detection algorithms, deep learning has higher accuracy and faster speed. At present, deep learning methods have been widely used in the field of fault diagnosis of insulators[4-8]. The detection scheme combined with deep learning can greatly improve work efficiency. But the mainstream solutions are all aimed at visible light datasets. The visible light dataset can reflect the external contour of the insulator well, but cannot reflect the internal condition of the insulator, which is not conducive to subsequent fault diagnosis. Infrared images can well reflect the heat radiated from the inside and use colors to distinguish the temperature difference between the power equipment and its background[9-10]. Common failures such as mechanical failures, unreasonable electrical load conditions and internal defects will cause internal temperature abnormalities, and further color differences can be used to determine the failure of the equipment. Therefore, infrared data is more suitable for fault diagnosis of insulators.

We proposed a target detection model based on the Centernet network[11]. We introduce the Centernet network into the insulator target detection to improve the detection accuracy of the algorithm
in the complex background. Our model can meet people's accuracy requirements for insulator target detection in actual application environments.

2. Centernet network principle

Today's mainstream algorithm models such as Yolov3, Faster-CNN, etc. need to list all possible frames of each category on the feature map, and each possible location will list multiple sizes and shapes of frames to cater to the shape of the target. This method is extremely computationally intensive and requires post-processing to remove redundant prediction frames. This leads to the overall complexity of the model and the inability to achieve end-to-end training. Centernet provides a unique detection idea. The Centernet model predicts the center point of the target, then the model directly predicts the width and height information of the target at the center, and finally generates a prediction box that conforms to the shape of the target. This method can greatly reduce the detection speed of the model, and at the same time can use the Maxpool layer to achieve end-to-end training.

The Centernet model will adjust the input picture in the initial stage, unify the input picture size to 512×512 through methods such as cropping, completion, etc., and downsample the input picture size to 128×128 in the preprocessing stage. Subsequently, the Centernet network provided a variety of backbone networks. The backbone network is the main body of the Centernet model, which is mainly used to perform feature learning on input image data and output feature maps. Finally, the feature map will be sent to three modules to predict the center point of the target, the floating point offset of the center point, and the width and height information. Figure 1 shows the overall structure of Centernet.

![Centernet structure diagram](image)

3. Loss Function

The overall loss function of the Centernet network is shown in formula (1), where size is 0.1 and off is 0.9.

$$L_{det} = L_k + L_{size} \lambda_{size} + L_{off} \lambda_{off}$$

(1)

3.1 offset loss

Because the center point of the Centernet network has a certain floating point error in the prediction process. Therefore, the Centernet network uses the L1 Loss function to train the offset value of the center point. The specific formula is shown in (2). Among them, $\hat{O}_p$ is the predicted bias value, and $(\frac{P}{R} - \hat{P})$ is the actual bias value of the network in the prediction process. The formula (2) reduces the floating point offset loss of the center point by comparing the L1 norm of the two, so it can realize more accurate prediction.
3.2 Center point prediction loss

The specific formula of the center point prediction loss function is shown in (3). Among them, $\alpha$ and $\beta$ are hyperparameters, which are set to 2 and 4 in the Centernet network. $N$ is the number of key points of the image.

\[
L_{off} = \frac{1}{N} \sum_p |\hat{P}_p - \frac{P}{R} - \hat{P}|
\]  

(2)

\[
L_{off} = -\frac{1}{N} \sum_{xyc} \left\{ \begin{array}{ll}
(1 - \hat{Y}_{xyc})^\alpha \log(\hat{Y}_{xyc}) & \text{if } Y_{xyc} = 1 \\
(1 - Y_{xyc})^\beta (\hat{Y}_{xyc}) \log(1 - \hat{Y}_{xyc}) & \text{else}
\end{array} \right. 
\]

(3)

3.3 Width and height prediction loss

The width and height data of the target $k$ and the corresponding center point coordinates are expanded according to formula (4). The Centernet network uses the value of $Y_{xyc}$ to predict the center point, and at the same time realizes the regression of the width and height dimensions of each target. Finally, the L1 function is used as the loss function of the width and height prediction. Equation (5) is the width and height loss function, where $\hat{S}$ is the predicted value, and $S_k$ is the length and width value after downsampling.

\[
p_k = \left( \frac{x_1^{(k)} + x_2^{(k)}}{2}, \frac{y_1^{(k)} + y_2^{(k)}}{2} \right)
\]

(4)

\[
L_{size} = \frac{1}{N} \sum_p |\hat{S}_{p_k} - S_k|
\]

(5)

4. Experiment procedure

In order to ensure the reliability of the experimental data, the experiments in this paper are all trained and tested on the laboratory server. The server configuration is shown in Table 1. We run the Centernet network model in the Pytorch1.2 environment, training 32 samples in each batch, and training a total of 140 batches. The learning rate was initially set to $1.25 \times 10^{-4}$, and was reduced to 0.1 times in the 90th and 120th batches respectively.

| Table 1. server configuration. |
|-------------------------------|
| SYSTEM                      | ubuntu16.04LTS               |
| CPU                         | Intel Xeon W-2145            |
| GPU                         | RTX 2080Ti                   |
| RAM                         | 16GB DDR4×4                  |

Figure 2 shows how the loss value of the Centerent network decreases with batches. It can be seen from Figure 2 that the Centerent network model has a good declining speed, and the loss value can be significantly reduced to about 2 at about 20 batches, and finally reaches the lowest value at about 140 batches, and convergence is completed. At this time, the prediction box of the model can fit well with the label box.
5. Result analysis
The speed detection standard of the deep learning detection model is the time it takes to detect a single image, in ms. Accuracy is measured by AP (average precision) value. The specific calculation formula is as formula (6). Among them, $P_r$ represents the ratio of the insulators successfully detected by the model to the total detection targets. $R_e$ represents the completeness of the model’s feature detection of insulators. $t_p$ (true positives) represents the number of insulators correctly detected by the model. $f_p$ (false positives) represents the number of insulators identified as background. $f_n$ (false negatives) represents the number of insulators recognized as background. Finally, when analyzing the results, we will calculate the detection accuracy of the insulators and use the AP value to evaluate the overall performance of the model.

$$P_r = \frac{t_p}{t_p + f_p}, \quad R_e = \frac{t_p}{t_p + f_n}$$  \hspace{1cm} (6)

In the original Centeret, the author provided a total of four backbone networks: DLA34, Res18, Res101 and Hourglass. We compare these four types of backbone networks, and the results are shown in Table 2.

Table 2. Comparison of detection results of different backbone networks.

| Insulator  | Speed |
|------------|-------|
| DLA34      | 0.932 | 25ms |
| Res18      | 0.949 | 12ms |
| Res101     | 0.924 | 23ms |
| Hourglass  | 0.857 | 54ms |

It can be seen from the data in the table that Res18 has the highest accuracy among the four backbone networks, which has an accuracy advantage of 2%-9% compared to the other three networks. At the same time, the speed of Res18 is the highest among the four backbone networks, which is very
conducive to the model's real-time detection of insulators. Therefore, Res18 is more suitable as the backbone network of the Centernet model.

We compared the Centernet model with Res18 as the backbone network with several mainstream detection models. As shown in Table 3, our network has a 4% to 5% improvement in accuracy compared to the three mainstream models, and at the same time performs best in terms of speed, slightly better than Yolov3. Figure 3 shows the Center point prediction map and the detection effect of our model on some datasets. On the whole, the overall performance of the Centernet model we proposed is due to several other mainstream models, which is very suitable for real-time fault diagnosis systems for power equipment.

Table 3. Comparison of detection effects of different models.

| Insulator   | Speed |
|-------------|-------|
| Yolov3      | 0.896 | 16ms |
| Faster-CNN  | 0.891 | 65ms |
| Centernet   | 0.949 | 12ms |

Figure 3. The Center point prediction map and the detection effect of our model.

6. Conclusion

The complex environment will affect the model's recognition rate of insulators. We introduce the Centernet network into insulator detection, and conduct training and testing on our own infrared data set. The results show that our proposed model has a significant improvement in accuracy of 4% to 5% compared with several mainstream models, and at the same time performs best in terms of speed. This will facilitate the real-time detection of infrared images of insulators in the next step.

Acknowledgments

This research was supported in part by the National Natural Science Foundation of China under Grant 51907034, in part by the Guangxi Science and Technology Base and Talent Project under Grant 2020AC19010.

References

[1] Zhao Z, Le W. Aerial insulator string image automatic location method[J]. Chinese Journal of Scientific Instrument, 2014, 35(3): 558-565.
[2] Reddy M J B, Mohanta D K. Condition monitoring of 11 kV distribution system insulators
incorporating complex imagery using combined DOST-SVM approach[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2013, 20(2): 664-674.

[3] Murthy V S, Gupta S, Mohanta D K. Digital image processing approach using combined wavelet hidden Markov model for well-being analysis of insulators[J]. IET Image Processing, 2011, 5(2): 171-183.

[4] Reddy M J B, Chandra B K, Mohanta D K. A DOST based approach for the condition monitoring of 11 kV distribution line insulators[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2011, 18(2): 588-595.

[5] Murthy V S, Gupta S, Mohanta D K. Distribution system insulator monitoring using video surveillance and support vector machines for complex background images[J]. International Journal of Power and Energy Conversion, 2009, 1(1): 49-72.

[6] Murthy V S, Mohanta D K, Gupta S. Power system insulator condition monitoring automation using mean shift tracker-FIS combined approach[J]. International Journal of Computer Aided Engineering and Technology, 2013, 5(1): 1-19.

[7] Murthy V S, Tarakanath K, Mohanta D K, Gupta S. Insulator condition analysis for overhead distribution lines using combined wavelet support vector machine (SVM)[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2010, 17(1): 89-99.

[8] HSIEH J W, HSU Y T, LIAO H Y M, et al. Video-based human movement analysis and its application to surveillance systems[J]. IEEE Transactions on Multimedia, 2008, 10(3): 372-384.

[9] Mohd Shawal Jadin, Soib Taib. Recent progress in diagnosing the reliability of electrical equipment by using infrared thermography[J]. Infrared Physics & Technology, 55(4).

[10] Jadin M S, Taib S, Ghazali K H. Finding region of interest in the infrared image of electrical installation[J]. Infrared Physics & Technology, 2015, 71: 329-338.

[11] DUAN K W, BAI S, XIE L X, et al. CenterNet: keypoint triplets for object detection[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South). IEEE, 2019: 6568-6577.