



**HAL**  
open science

# Deep Learning-Based Segmentation of Key Objects of Transmission Lines

Mingjie Liu, Yongteng Li, Xiao Wang, Renwei Tu, Zhongjie Zhu

► **To cite this version:**

Mingjie Liu, Yongteng Li, Xiao Wang, Renwei Tu, Zhongjie Zhu. Deep Learning-Based Segmentation of Key Objects of Transmission Lines. 19th International Conference on Entertainment Computing (ICEC), Nov 2020, Xi'an, China. pp.317-324, 10.1007/978-3-030-65736-9\_29 . hal-03686012

**HAL Id: hal-03686012**

**<https://inria.hal.science/hal-03686012>**

Submitted on 2 Jun 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Deep Learning-Based Segmentation of Key Objects of Transmission Lines

Liu Mingjie,<sup>1</sup> Li Yongteng,<sup>1</sup> Wang Xiao,<sup>1</sup> Tu Renwei,<sup>2</sup> and Zhu Zhongjie<sup>2,\*</sup>

<sup>1</sup> Ninghai Power Supply Company Limited, State Grid Corporation of Zhejiang, 315600 Ningbo, China

<sup>2</sup> Zhejiang Wanli University, College of Information and Intelligence Engineering, 315100 Ningbo, China  
zhongjiezhu@hotmail.com

**Abstract.** UAV(Unmanned Aerial Vehicle) inspection is one of the main ways of transmission line inspection, which plays an important role in ensuring the transmission safety. In view of the disadvantages of existing inspection methods, such as slow detection speed, large calculation of detection model, and inability to adapt to low light environment, an improved algorithm based on YOLO(You Only Look Once) v3 is proposed to realize the real-time detection of power towers and insulators. First of all, a data set of power towers and insulators is established, which are inverted and transformed to expand the data volume. Secondly, the network structure of YOLO v3 is simplified and the calculation of the detection model is reduced. Res unit is added to reuse convolution feature. Then, K-means is used to cluster the new data set to get more accurate anchor value, which improves the detection accuracy. Through the experimental demonstration, the accuracy of the proposed scheme for the detection of key parts of the transmission line is 4% higher than the original YOLO, and the detection speed reaches 33.6 ms/frame.

**Keywords:** YOLO v3, Transmission Line, Tower, Insulator.

## 1 Introduction

The safe and stable operation of transmission lines is of great significance for the regional social and economic development. In order to eliminate the potential safety hazards before power failure, it is necessary to carry out daily inspection of transmission lines [1]. The traditional inspection method relies on manual inspection, which has great potential safety hazards. Especially after the natural disasters such as earthquake and typhoon, the inspection environment becomes worse and more complex. Moreover, the manual inspection is time-consuming, inefficient, labor intensive and cost-effective. The UAV(Unmanned Aerial Vehicle) inspection is a kind of inspection method based on machine learning and target detection algorithm, which can be realized without directly contacting the power system. Compared with manual inspection, UAV inspection has the advantages of low risk, high efficiency and low cost

[2][3][4]. Therefore, UAV patrol technology is widely concerned by the power industry.

In traditional machine vision recognition and location methods, based on shallow machine learning algorithms such as SIFT (Scale Invariant Feature Transform) recognition method, combined with HOG (Histogram of Oriented Gradients) and SVM (Support Vector Machine) and other methods are mainly based on the edge or texture features in the image. This part of the algorithm is relatively simple in the background environment, and has achieved good results in the image with clear edge of the tested object. In the image scene recognition, T. Yu and R. Wang put forward the algorithm of scene analysis using image matching, and achieved good results in the street view. However, the transmission line image captured by UAV contains a lot of environmental background information [5]. In the face of complex backgrounds, traditional algorithms cannot meet the real-time and accuracy of detection.

In recent years, deep learning, especially convolution neural network model, has achieved remarkable results in image recognition. R. Girshick et al. Proposed RCNN [6], which changed the traditional region selection to use sliding window. Each sliding window is detected once, RCNN uses a heuristic method (selective search), and then detects candidate regions, reducing the degree of information redundancy. However, the image scale fixed in CNN will cause the object deformation and then lead to the degradation of detection performance. Later, S. Ren proposed fast RCNN [8], which really realized the end-to-end target detection framework and improved the detection accuracy and speed. Then, Redmon and others proposed the YOLO algorithm [9], which greatly improved the speed of target detection. Among them, the ability of YOLO v3 to detect small targets has been improved obviously, which makes its performance to detect different sizes of targets more balanced. In the detection of transmission towers and insulators, YOLO v3 mainly detects large and medium-sized targets, but almost no small targets. There is no doubt that the part of small target detection will increase the calculation of the whole detection process. At the same time, in the low light environment, there is such a situation of missing or false detection.

To resolve the abovementioned problems, this paper improves upon the YOLO v3 framework to realize real time object detection of key parts of the transmission line. In order to make up for the loss of features caused by network simplification, res unit is added after the backbone network to reuse the convolution features [10]. Since there is no public data set for key parts of transmission lines, a new data set with weak light conditions is established. The data set comes from UAV multi angle shooting, which simulates the real detection environment. Also, it is rotated and scaled to expand the data. According to the characteristics of the developed data set, the optimal clustering center point is selected and reclustered using the K-means approach to obtain more accurate anchors.

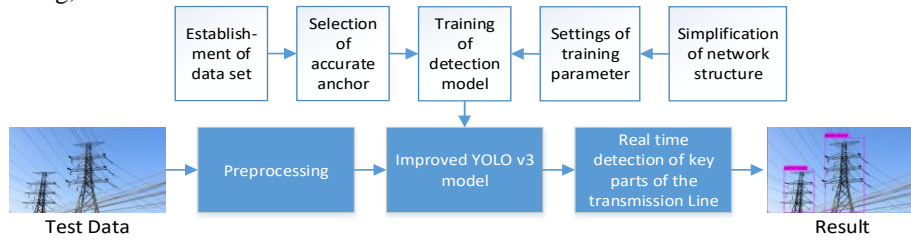
The contributions of this work can be summarized as follows: (1) Establishment of a data set [11]. (2) Simplification of network structure. (3) Selection of accurate anchors.

In Section 2, we introduce the new data set, propose the optimization scheme of the network structure of YOLO v3, and select a more accurate anchor value. Section 3

shows the experimental results and analysis. Finally, section 4 draws the conclusion of this paper.

## 2 Proposed scheme

Based on the improved YOLO v3, this paper proposes a real-time location detection method for key parts of transmission line. By building more targeted data set of power tower and insulator, and selecting more accurate anchor value through K-means clustering, the real-time detection of transmission towers and insulators can be realized.



**Fig. 1.** Diagram of the proposed scheme

The flow chart of the detection method in this paper is shown in Figure 1. For two types of detection targets of key parts of transmission line: towers and insulators, a new data set is established. Lighting, background and shooting angles are fully considered. After that, two kinds of detection targets in the data set are labeled with YOLO-Mark. And then more accurate anchor value are obtained through K-means clustering. The data set is put into the optimized convolution neural network to train the detection model. Finally, the results can be obtained by inputting the test data into the improved YOLO v3 model.

### 2.1 Establishment of data set

The training data set of neural network usually require at least a few thousand. If the picture is too few, it will affect the accuracy and reliability of detection. Because there is no public data set of the key parts of the transmission line, a new data set is established by collecting online pictures and UAV photos. The pictures are expanded by the way of inversion and scale transformation.

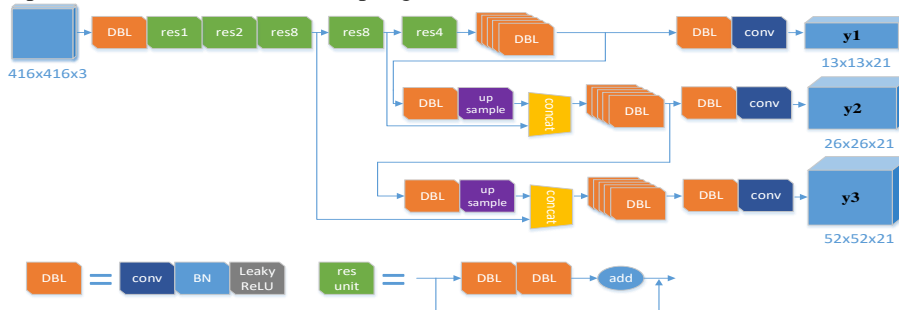
These images contain different illumination, different shooting angles of detection objects, different resolutions, different detection backgrounds and other conditions, which meet the requirements of sample diversity, and make them get purposeful optimization. The equalization of samples is of great significance to improve the robustness of the algorithm. If the illumination factor is not taken into account in the new data set, only a single illumination sample will lead to obvious false detection or missing detection when the trained model detects dark pictures or videos. Scientific data set can achieve better detection results with fewer training samples. Sample data set is as follows:



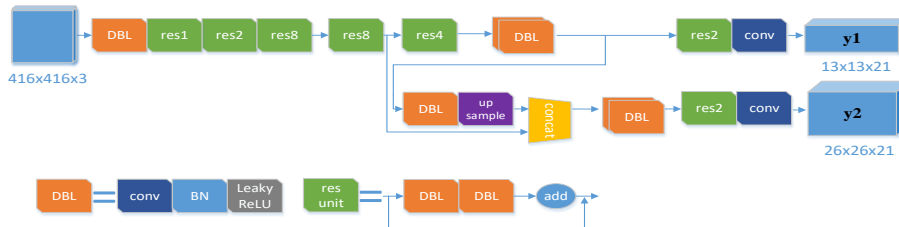
**Fig. 2.** Sample images in the developed data set

## 2.2 Simplification of network structure

YOLO v3 is proposed by Joseph Redmon and Ali Farhadi of Washington University in the United States. Its detection speed is very fast and its detection accuracy is high. In the aspect of basic image feature extraction, YOLO v3 adopts a network structure called darknet-53 [12][13], which has 53 convolution layers in total. In order to better extract features, some parts of the structure refer to the method of residual network and set up quick links between some layers. In the whole structure, since there is no pooling layer and full connection layer, when the convolution layer is transferred, the tensor size transformation is realized by changing the convolution kernel moving step length, such as  $\text{strip} = (2,2)$ , which is equal to dividing the image side length by 2, making the area  $1/4$  of the original [14]. In the whole convolution process, YOLO v3 experienced 5 times of down sampling.



(a) YOLO v3



(b) Improved YOLO v3

**Fig. 3.** Comparison of the network structure of the original and improved YOLO v3 frameworks

In order to improve the detection ability of small targets, the network structure of YOLO v3 is optimized. YOLO v3 realizes feature fusion by concatenating convolution layers of different depths and depths, and constitutes FPN structure by three sets of feature graphs of different sizes. The feature outputs of  $13 \times 13$ ,  $26 \times 26$  and  $52 \times 52$  are respectively for the detection of large, medium and small targets. In this paper, the feature output of  $52 \times 52$  for small target detection is eliminated. In addition, res unit is added behind the backbone network to enhance features. Before the output of  $y_1$  and  $y_2$ , four DBLs are replaced by two res units, which the size and number of convolution cores are kept unchanged.

### 2.3 Selection of accurate anchor

K-means algorithm is a typical distance based clustering algorithm, which evaluates the similarity according to the distance, that is, the closer the distance between two objects, the greater the similarity. Clusters are made up of objects close to each other, so the final goal is to get compact and independent clusters. In this paper, K-means is used to calculate the anchor value, because using Euclidean distance will cause more errors in large bounding boxes than in small ones. And then get good IoU scores through anchors.

IoU is the ratio of intersection and union of prediction box and annotation box. The larger the ratio is, the better the detection performance of the detection model is, and the higher the accuracy is. In this paper, 1 to 9 clustering centers are set for the data set, and the anchor and IoU values obtained from clustering are shown in the following table:

**Table 1.** Results of K-means clustering.

K	Anchor	Avg IoU
1	54, 118	21.21%
2	30, 61, 138, 324	38.31%
3	26, 47, 63, 182, 173, 364	43.83%
4	24, 42, 57, 129, 95, 357, 232, 353	47.19%
5	29, 24, 13, 56, 84, 63, 34, 167, 153, 356	58.53%
6	12, 48, 33, 27, 30, 145, 92, 67, 91, 344, 226, 359	59.58%
7	12, 50, 28, 22, 59, 50, 30, 151, 132, 89, 88, 348, 217, 368	60.15%
8	11, 45, 29, 23, 24, 109, 63, 50, 43, 215, 136, 91, 101, 367, 234, 364	61.06%
9	11, 44, 27, 22, 22, 90, 58, 44, 34, 182, 116, 84, 69, 328, 133, 374, 267, 349	61.88%

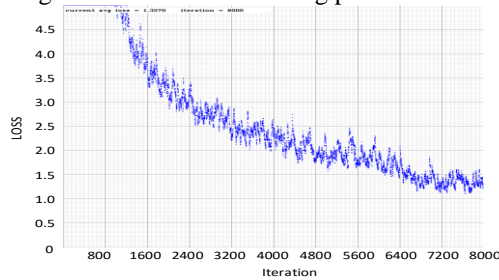
The above table shows different anchor values and different Avg IoU obtained by setting different number of cluster centers. With the increase of cluster center, Avg IoU shows an increasing trend, the increasing speed is fast first and then slow, and gradually tends to convergence. When  $k = 6$ , Avg IoU reached 59.58%, and then became stable. Therefore, in order to reduce the amount of calculation as much as pos-

sible, this paper selects the anchor value when  $k = 6$ : (12, 48), (33, 27), (30, 145), (92, 67), (91, 344), (226, 359).

### 3 Experiments

The experiment development environment is as follows: CPU: Intel i7. GPU: NVIDIA GeForce GTX1060. RAM: 16 G. Deep learning network framework: Darknet-53.

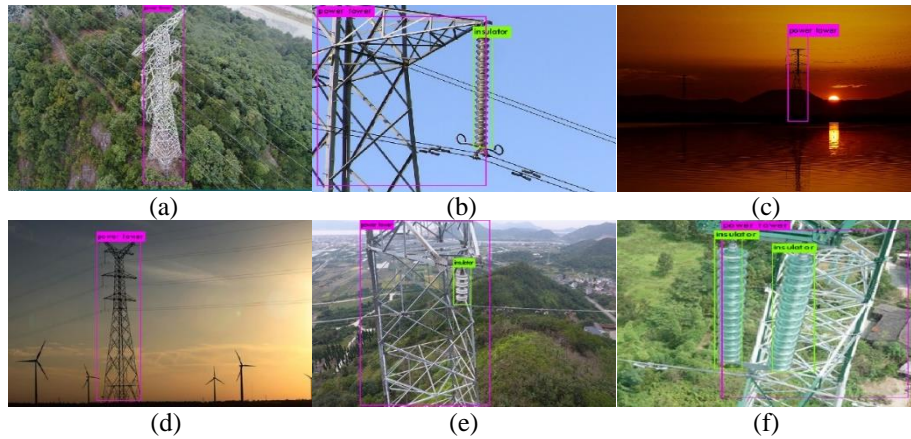
In order to train the model better and get better detection results, the training parameters of YOLO v3 are optimized. In this paper, there are two kinds of detection targets, transmission tower and insulator. Each type of detection target has 4000 iterations. The input resolution is  $416 \times 416$ , and multi-scale training is started. The learning rate determines the speed of weight updating. If it is set too much, the result will exceed the optimal value, and if it is too small, the descent speed will be too slow. So a dynamic learning rate is set up to get a better model. When setting  $0 < \text{iteration} < 6400$ ,  $\text{lr} = 0.001$ ;  $6400 < \text{iteration} < 7200$ ,  $\text{lr} = 0.0001$ ;  $7200 < \text{iteration} < 8000$ ,  $\text{lr} = 0.00001$ , and the learning rate of the whole training process decreases by 100 times.



**Fig. 4.** Loss curve

The loss curve above is obtained through training. The early training  $\text{lr} = 0.001$  is set to make the loss rapidly decrease. From the 1600th iteration, the speed of loss decrease slowly becomes more and more stable. From the 6400th iteration, it tends to converge. Finally, the loss value converges to 1.3270. The ideal small sample training result is obtained.

The reliability of the model is verified by a large number of picture tests. The experimental results are as follows: The detection accuracy of the optimized YOLO v3 for transmission towers and insulators is quite reliable, and which meets the requirements of real-time detection. Compared with the existing inspection method, the algorithm in this paper can not only ensure the detection accuracy, but also reduce the amount of calculation and improve the detection speed. The new data set can still have high detection performance in low light conditions, especially in extreme weather or natural disasters before and after the scene is still able to perform inspection tasks.



**Fig. 5.** Partial test results

Six representative experimental results are selected for analysis, (a) (b) (e) (f) represent the results of single target and multi-target detection in complex background. Because of the diversity of data sets, the detection model has the ability to detect the transmission tower with partial structure, and also reflects the generalization ability of the model. The improvement of convolutional neural network greatly improves the ability of the trained model to detect large and medium size targets. (c) and (d) are the detection in low light environment. Single target and multi-target in different background can also be detected accurately.

**Table 2.** Comparison of test results.

Algorithm	Recall	Average IoI	Precision	Speed	Total BFLOPs
YOLO v3	0.42	62.06%	81%	33.2 ms/frame	65.297
Improved YOLO v3	0.37	62.66%	85%	33.6 ms/frame	63.097

Table 2 shows the data comparison between original YOLO v3 and the improved YOLO v3. The recall and average IoU of the improved YOLO v3 are almost the same as the original. At the same time, precision is increased by 4%, Total BFLOPs is reduced by 2.2, and real-time detection speed is guaranteed.

## 4 Conclusion

Based on the improved YOLO v3, an algorithm to detect the key parts of transmission line is proposed. The data set of transmission tower and insulator are newly established, and pictures in low light environment are added. According to the detection requirements of large and medium-sized targets, a simplified network structure of YOLO v3 is proposed to reduce the calculation of the model. At the same time, res unit is added behind the backbone network to make feature reuse and make up for the



loss of feature brought by structure simplification. Furthermore, K-means clustering is used to get more accurate anchor value, which makes the detection accuracy further improved. In addition, the experimental results show that the detection accuracy of the improved YOLO v3 model is 85%, and the detection speed is 33.6ms/frame, which has a certain detection ability under low light conditions. It will improve the efficiency of UAV inspection in the actual scene, and has high application value.

## Acknowledgment

This work was supported in part by the National Natural Science Foundations of China (No. 61671412); Zhejiang Provincial Natural Science Foundation of China (No. LY19F010002); the Natural Science Foundation of Ningbo, China (2018A610053); Ningbo Municipal Projects for Leading and Top Talents (NBLJ201801006); and Innovation and consulting project from Ninghai Power Supply Company, State Grid Corporation of Zhejiang, China.

## References

1. Zormpas A., "Power Transmission Lines Inspection using Properly Equipped Unmanned Aerial Vehicle (UAV)," *IEEE Instrumentation and Measurement Society*, pp. 1-5 (2018).
2. Miao X., "Insulator Detection in Aerial Images for Transmission Line Inspection Using Single Shot Multibox Detector," *IEEE Access*, vol. 7, pp. 9945-9956 (2019).
3. Yang Q., "A New Fusion Chemical Reaction Optimization Algorithm Based on Random Molecules for Multi-Rotor UAV Path Planning in Transmission Line Inspection," *Journal of Shanghai Jiaotong University (Science)*, vol. 23, pp. 671-677 (2018).
4. Bhola R., "Detection of the power lines in UAV remote sensed images using spectral-spatial methods," *Journal of Environmental Management*, vol. 206, pp. 1233-1242 (2018).
5. Wang Y., "Detection and Recognition for Fault Insulator Based on Deep Learning," pp. 1-6 (2018).
6. Zhao Z., "Insulator Detection Method in Inspection Image Based on Improved Faster R-CNN," *Energies*, vol. 12, p. 1204 (2019).
7. Sengupta, Abhronil, et al. "Going Deeper in Spiking Neural Networks: VGG and Residual Architectures." *Frontiers in Neuroence* (2018).
8. Qiao Y., "Cattle segmentation and contour extraction based on Mask R-CNN for precision livestock farming," *Computers and Electronics in Agriculture*, vol. 165, p. 104958, 2019.
9. Lu J., "A Vehicle Detection Method for Aerial Image Based on YOLO," *Journal of Computer and Communications*, vol. 06, pp. 98-107 (2018).
10. Leng J. and Liu Y., "An enhanced SSD with feature fusion and visual reasoning for object detection," *Neural Computing and Applications* (2018).
11. Zhao W., "Research on the deep learning of the small sample data based on transfer learning," *AIP Conference Proceedings*, vol. 1864 (2017).
12. Redmon J. and Farhadi A., "YOLOv3: An Incremental Improvement," (2018).
13. Ju, "The Application of Improved YOLO V3 in Multi-Scale Target Detection," *Applied Sciences*, vol. 9, p. 3775, 2019.
14. Kim K., "Performance Enhancement of YOLOv3 by Adding Prediction Layers with Spatial Pyramid Pooling for Vehicle Detection," pp. 1-6 (2018).