

Power Tower Tilt Monitoring Method Based on Beidou and Improved Yolov3

Lei Sun^{1*}, Caiguo Ma¹, Hanqin Hu², Tao Ni², Mingquan Zhang³, Zeen Zhang³

¹Hangzhou Kaida Electricity Construction Co., Ltd. Installation Undertaking Branch Office, Hangzhou, China

²State Grid Zhejiang Electric Power Co., Ltd. Hangzhou Power Supply Company Yuhang Branch, Hangzhou, China

³North China Electric Power University, Baoding, China

Abstract

Aiming at the problems of low efficiency, long patrol cycle, and poor accuracy of deformation judgment in existing inspection methods of power poles and towers, this paper proposes a UAV-assisted tower rod tilt monitoring method based on BeiDou satellite navigation, and introduces the tower identification algorithm and tilt detection algorithm in detail. For tower recognition, CFA and improved DarkNet53 were introduced on the basis of the original YOLOv3, and the F value reached 85.1%, which significantly improved the detection accuracy of small targets such as tower edges. For tilt detection, Canny edge detection algorithm and LSD straight line detection method were used to extract the center line and determine the tilt Angle, and the final accuracy reached 87.4%.

Keywords

Power tower tilt, YOLOv3, Canny, LSD, BeiDou

1 Introduction

There are many kinds of geological disasters such as landslides in mountainous areas due to the harsh environmental conditions, complex geological topography and changeable climate around the transmission tower site. Frequent geological disasters are easy to lead to the tilt, displacement, deformation, settlement of the transmission line tower or foundation and surrounding landslides, which pose a serious threat to the safe operation of the power grid. With the development of new technology, tower tilt inspection has gradually developed from manual inspection to robot and UAV inspection^[1].

The Beidou Navigation Satellite System (BDS) is a global navigation and positioning system developed by China. It has the capability of navigation, precise positioning and timing, and the Beidou 3 system will be fully completed by 2020 to provide global users with services such as timing, positioning and navigation, global short message communication and international search and rescue^[2]. The application of Beidou technology to real-time monitoring of health status of high-voltage transmission towers is a successful application of "Internet + space-based information System" in the field of smart grid. Beidou satellite has the advantages of wide observation range and high efficiency, and can realize all-day and all-weather monitoring, which provides a new method for power tower tilt monitoring and fault discovery.

In this paper based on beidou and improve YOLOv3 power tower tilting monitoring method of tower recognition algorithm and gradient detection algorithm was introduced in detail the implementation of the scheme and process, and the algorithm is verified by experiment can achieve high accuracy of real-time monitoring of transmission line tower, is of great significance to ensure the trouble-free operation of tower.

AIoTC2022@International Conference on Artificial Intelligence, Internet of Things and Cloud Computing Technology

EMAIL: *346245211@qq.com (Sun Lei)



© 2022 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

2 Related work

2.1 Object detection based on deep learning

In 2015, Girshick and Ren et al. proposed Fast R-CNN^[3] and Faster R-CNN^[4], which greatly improved the detection speed, the rate could reach 5FPS, and the mAP on VOC2007 dataset was increased to 68%. In 2016, Redmon et al. proposed a target detection method YOLO(You Only Look Once) with high detection speed and high accuracy^[5], and the real-time rate can reach 45 frames per second. Liu et al. proposed the SSD(Single Shot Multi Box Detector)^[6] algorithm, which combines the characteristics of YOLO and Faster R-CNN, but the detection effect of overlapping objects and small scale objects is poor. In 2017, Lin et al. proposed Feature Pyramid Networks (FPN)^[7], which firstly generated multi-scale Feature maps from the bottom up, and then fused low-resolution Feature maps with high-resolution Feature maps by means of top-down paths and lateral connections. This structure can make full use of the semantic information of each feature map of different scales, and effectively improve the accuracy of small object detection. In 2018, Redmon proposed YOLOv3^[8] algorithm, adding a multi-scale recognition module on the original basis, which not only ensured the speed, but also improved the recognition ability of small targets, and reached an average accuracy of 33% on the MSCOCO dataset. Liu et al. proposed an improved Path Aggregation Network (PANet) for Instance Segmentation^[9] based on FPN, which added a bottom-up Path enhancement structure on the basis of FPN to better preserve the shallow feature information. Finally, a fully connected layer is used to capture different views of each candidate region to achieve better prediction results. In 2021, Hu et al. proposed the Micro-YOLO^[10] algorithm based on YOLOv3, replacing the original convolution layer with depthseparable convolution, reducing the number of parameters by 3.46 times and the product operation by 2.55 times.

2.2 Tilt detection of power tower

Shen et al. reconstructed the tower point cloud model by ground 3D LiDAR, and then measured the tower tilt of transmission lines by manual comparison, but the measurement results were directly affected by the accuracy of the point cloud model and subjective factors^[11]. Zhang et al. designed a real-time monitoring system for transmission tower tilt, which uses low-power wide area network technology for data transmission, and the monitoring effect largely depends on the communication quality^[12]. In addition, the commonly used methods of tower toppling detection also include plumb method, plane mirror method, theodolite method and sensor method, but these methods generally have the problems of high requirements for operators, large amount of work, high risk and low efficiency.

In recent years, with the maturity of UAV technology and the development of deep learning technology, a major breakthrough has been made in the working mode of collecting tower images with UAV and other tools and then sending them back to the platform for analysis using object detection technology. Lu et al. calculated the center line of the tower by fitting the tower structure to obtain the tilt rate, but the accuracy of the point cloud model directly affected the measurement results. Lei et al.^[13] successfully applied the Faster-RCNN algorithm to the abnormal state detection of transmission lines, but the algorithm required high computational power and long analysis time, so it was difficult to apply to edge devices. Guo et al.^[14] improved the detection speed by simplifying the YOLO algorithm to realize real-time detection of the tower, which provided a good technical idea. However, this model can only roughly classify and locate the state of the tower, and cannot calculate the tilt Angle of the tower. Moreover, the classification accuracy is very dependent on the training set, and the generalization ability of the model is poor.

3 The method in this paper

3.1 The main process

In daily situations, when high power facilities such as poles and towers continue to tilt or twist 1-2cm in one direction every day, there will be a risk of tower collapse. The research shows that the ad-

ditional loss caused by power system failure is more than 400 times of its own failure loss. At present, the route patrol application of unmanned aerial vehicle (UAV) mainly uses the mobile performance of UAVs as an auxiliary acquisition tool of field route images to reduce the difficulty of field information collection. However, the evaluation of field conditions such as the status of line poles and towers still needs to rely on manual judgment. The functions of precision timing, high-precision measurement, positioning and navigation and short message communication of the Beidou system effectively solve the problem that UAV requires personnel to control, resulting in inspection efficiency and quality mainly determined by human hands.

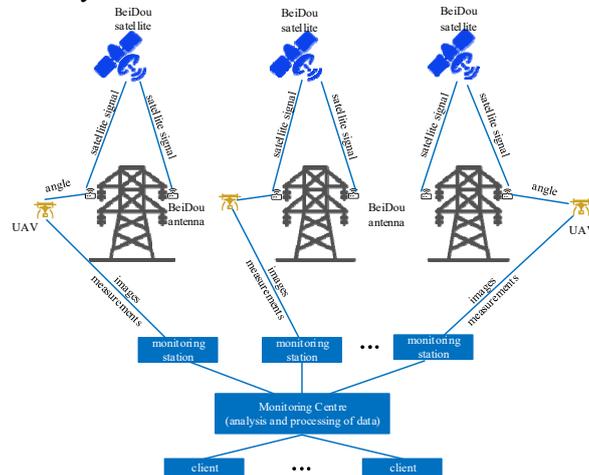


Figure 1 System construction

The tilt monitoring structure of power tower based on BeiDou and improved YOLOv3 proposed in this paper is shown in **Figure 1**. First BeiDou three generations of attitude measuring technology is used for real-time monitoring of tower tilt, in particular, by receiving satellite, BeiDou satellite signals under the epoch, build a baseline of carrier phase double difference equation L attitude measuring mathematical model, and use the SPSA algorithm rapid baseline L attitude Angle, then tilt data obtained. Next, the BeiDou antenna sends the collected tower tilt data to the corresponding UAV, which determines the shooting Angle according to the tower pole position and tilt data. Finally, the BeiDou short message or 4G and other wireless communication technologies are used to send the aerial image of tower rod and tilt data to the monitoring center and send them to the cloud monitoring platform. According to the tower rod image, the cloud monitoring platform firstly uses the improved YOLOv3 technology to detect the main body of tower rod. Subsequently, Canny edge detection algorithm and LSD straight line detection method are used to extract the center line and determine the tilt Angle of the detected tower rod body. The trend of deformation of the tower is obtained through analysis, which provides a basis for relevant departments. The system can detect the hidden danger of the tower safety accident as soon as possible, and notify the relevant personnel of power to eliminate the hidden danger in time, so as to effectively avoid the occurrence of the tower tilt, collapse, line breaking and other hazardous accidents.

3.2 Tower identification algorithm

Feature-based object detection methods are difficult to be applied in actual scenes due to their poor environmental adaptability and high background requirements. With the development of deep neural networks and the proposal of various convolutional neural networks in recent years, their environmental adaptability and accuracy have been able to meet almost all applications in complex scenes. Existing deep neural network frameworks for object detection include Fast-RCNN, Faster-RCNN, SSD, YOLO-v3, etc. Considering that the tower is a large target, the improved YOLOv3 neural network framework is selected. Compared with Faster-RCNN, the detection accuracy of small targets is slightly lower, but the number of network layers is lower. Compared with Faster-RCNN, it has Faster processing speed under the same equipment, which is suitable for practical application of UAV inspection.

YOLO is an end-to-end network, which rejects the idea of Faster-RCNN sliding window and completes the output from the original image input to the object position and category. YOLOv3 network is a fully convolutional network, which uses a large number of layer hopping connections of residuals. In order to reduce the negative gradient effect caused by pooling, the pooling layer is directly discarded and convolution is used for downsampling. In the network structure, 1×1 and 3×3 convolutional layers are mainly used to achieve the function of unifying the number of channels and compression features, respectively.

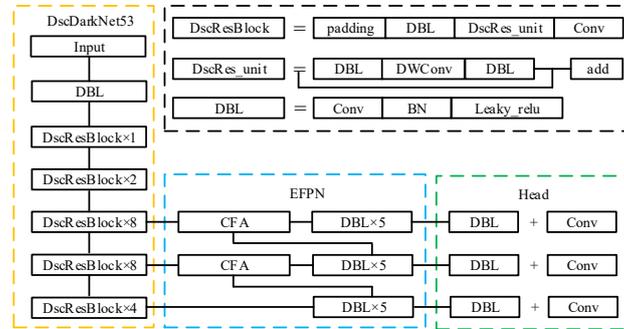


Figure 2 Improved YOLOv3 model

In view of the low detection accuracy of YOLOv3 small targets, this paper made two improvements on the YOLOv3 network, the structure of which is shown in **Figure 2**. (1) The original DarkNet53 network is introduced with Depth Separable Convolution (DSC)^[15]. DSC reduces the problem of target information loss caused by the deepening of the network. DSC can separate each channel of the input feature map and realize convolution operation respectively. Finally, point convolution is used to realize the combination of separate channel convolution and achieve the same effect as traditional convolution calculation, as shown in **Figure 3** (2) Cascading Fusion Attention (CFA) module is introduced between the layers of the original feature pyramid structure, as shown in **Figure 4**, to fuse the low-scale feature map with the corresponding high-scale feature map. Specifically, CFA can highlight local small sample information while emphasizing global features, so as to improve the detection ability of the model under extreme scale changes.

3.2.1 DscResBlock module

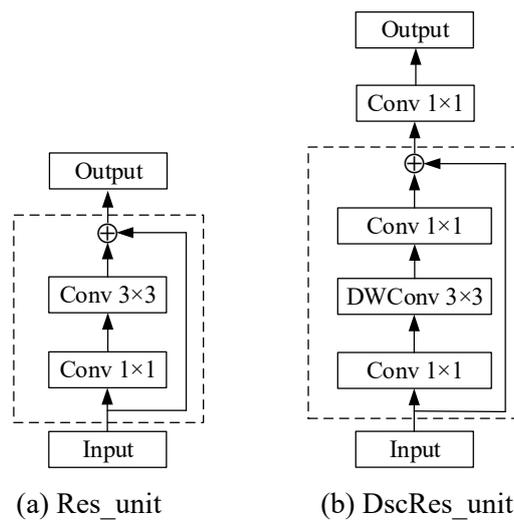


Figure 3 Residual structure of the original network and the improved residual structure

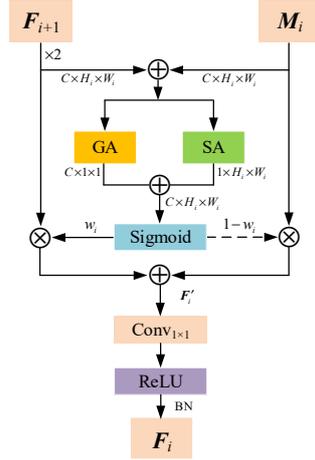


Figure 4 CFA Module Structure

Figure 3(a) shows the residual block of DarkNet53 network. Two different convolutions are directly used for feature extraction, and then jump connection is used to obtain the output. The structure designed in this paper is shown in **Figure 3(b)**. Firstly, 1×1 convolution is used for processing, and then 3×3 depth-separable convolution is added to perform convolution operation for channels and 1×1 convolution to fuse the separated channel information. Finally, 1×1 convolution is used for smooth output after jumping connection. DSC is introduced into all residual blocks in DarkNet53 to obtain DscDarkNet53, which can enhance the information interaction between channels, reduce the loss of feature information in the small area at the edge of the tower bar in the downsampling process, strengthen the feature information of the target to be detected, and facilitate the separation of the target from the complex background.

3.2.2 CFA module

The CFA module is composed of two modules, Global Attention Model (GA) and Spatial Attention Model (SA), whose input is two high-and low-scale feature maps respectively $M_i(i=1,2)$ and $F_{i+1}(i=1,2)$, the output is the fused feature map F_i . The specific steps are as follows:

- (1) Firstly, F_{i+1} was upsampled twice to unify the scale with M_i , and then the two were added pixel-by-pixel as the initial integration. The results were used to extract the spatial details of high and low scales through GA and SA, and the fusion weight with the same size as the feature map was generated, as shown in Equation (1).

$$w_i = \sigma(\text{GA}(M_i + \text{Up}_2(F_{i+1})) + \text{SA}(M_i + \text{Up}_2(F_{i+1}))), i = 1, 2 \quad (1)$$

Where, σ represents Sigmoid activation function, Up_2 represents double upsampling processing, GA and SA respectively represent global attention module and spatial attention module, $w_i(i = 1, 2)$ represents the fusion weight obtained by GA and SA processing.

- (2) Then, the fusion weight w_i obtained in the first step is used to dynamically select M_i and F_{i+1} at the element level, as shown in Equation (2).

$$F'_i = \text{Up}_2(F_{i+1}) \times w_i + M_i \times (1 - w_i), i = 1, 2 \quad (2)$$

The dashed line in **Figure 4** represents $1 - w_i$. The fusion weight w_i is composed of real numbers between 0 and 1. By combining with $1 - w_i$, the network can carry out weighted average between M_i and F_{i+1} , and the weighted feature map $F'_i(i = 1, 2)$ is output.

- (3) Finally, the fusion feature map is obtained by 1×1 convolution processing, as shown in Equation (3)

$$F_i = \delta(\text{Conv}_{1 \times 1}(F'_i)), i = 1, 2 \quad (3)$$

Where, δ represents the ReLU activation function, and then F_i is obtained by 1×1 convolution processing.

3.3 Tilt detection algorithm

The process of tilt detection in this paper is firstly to smooth and filter the image, and then to suppress the gradient amplitude with non-maximum value, so as to facilitate subsequent edge detection. Subsequently, Canny edge detection algorithm and LSD line detection method are used to extract the center line and determine the tilt Angle.

- (1) For image smoothing filtering, one-dimensional Gaussian function is used to smooth filter and denoise the image to be edge detected by row and column respectively, as shown in Equation (4).

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (4)$$

- (2) The gradient amplitude and direction are calculated, and the gradient amplitude and direction are obtained by finding the partial derivatives. The finite difference of the first partial derivatives in the 2×2 neighborhood is used to smooth and filter the first partial derivatives of the image (x,y) as shown in Equation (5).

$$\begin{cases} P_x(x,y) = [f(x,y+1) - f(x,y)]/2 \\ \quad + [f(x+1,y+1) - f(x+1,y)]/2 \\ P_y(x,y) = [f(x,y) - f(x+1,y)]/2 \\ \quad + [f(x,y+1) - f(x+1,y+1)]/2 \end{cases} \quad (5)$$

- (3) The edge localization can be more accurate by suppressing the gradient amplitude with non-maximum value. After refinement, the edge position can be determined by a single pixel. Specifically, the amplitude of the center pixel is compared with the two neighboring pixels around it. If the amplitude is greater than, the point is an edge point. If less than, the point is not an edge point.
- (4) The Canny algorithm uses double threshold T_h and T_l to segment the images with non-maximum suppressed values. If (x,y) is less than T_l , then the point must not be an edge point. If the gradient amplitude of (x,y) is greater than T_h , the point must be an edge point. If it is between the two, and $T_l < P(x,y) < T_h$, find any point near (x,y) that is greater than T_h . If so, it is an edge point, otherwise it is not an edge point.
- (5) Because the cross structure of line segments inside the tower will lead to the discontinuity of line segments at the edge when using the LSD algorithm for detection, it is necessary to connect multiple line segments in the same direction. Specific operations are as follows:

First according to the inclination of the filtered line will line group, grouping first will segment according to Angle size sorting, set two line Angle φ_i, φ_j respectively, such as Equation(6), (7), then judge after each line and line segment Angle difference is greater than 3° , if not greater than that of two line segments similar to belong to the same direction, Otherwise, separate the preceding segment from the following segment.

Secondly, each group of line segments can determine whether to fuse two adjacent line segments according to their own lengths and the shortest distance between adjacent lines, as shown in Equation (8), and the distance between the midpoint of the line segment and the line where the previous line segment is located, as shown in Equation (9). **Figure 5** shows the detection result.

$$\varphi_i = \arctan\left(\frac{y_{i2} - y_{i1}}{x_{i2} - x_{i1}}\right) \quad (6)$$

$$\varphi_j = \arctan\left(\frac{y_{j2} - y_{j1}}{x_{j2} - x_{j1}}\right) \quad (7)$$

$$l_{i,j} = \sqrt{\left(\frac{y_{i2} + y_{i1}}{2} - \frac{y_{j2} + y_{j1}}{2}\right)^2 + \left(\frac{x_{i2} + x_{i1}}{2} - \frac{x_{j2} + x_{j1}}{2}\right)^2} \quad (8)$$

$$d_{i,j} = \frac{(y_{j2} - y_{j1}) \left(\frac{x_{i1} + x_{i2}}{2} \right) + (x_{i1} - x_{i2}) \left(\frac{y_{i2} + y_{i1}}{2} \right) + x_{j2}y_{j1} - x_{j1}y_{j2}}{\sqrt{(y_{j2} - y_{j1})^2 + (x_{i2} - x_{i1})^2}} \quad (9)$$

- (6) In order to extract the contour of the tower, 2-4 longest line segments are first selected as candidates for the contour of the tower. Then selected slope angle is $\theta_i (0 < \theta_i < 90^\circ)$ respectively and $\theta_j (90^\circ < \theta_j < 180^\circ)$ of two line segments, judge $|\theta_i - (180^\circ - \theta_j)| < 5^\circ$ is established; If so, continue to select the longest group between two line segments from the pairwise combination as the tower contour; If not, pairwise combination will continue to find the line segment that meets the requirements. **Figure 5** shows the leaning angle of the tower.
- (7) Based on tower contour detection results, in order to get the tilt Angle, both sides need to compute the center line of contour, in particular, were selected two line upper endpoints y coordinates, the smaller the endpoint y coordinate larger points, respectively as A_1, B_1 as shown in **Figure 6**, horizontal parallel lines, and another line segment to the corresponding two A_2, B_2 , The midpoint M_1 of A_1B_1 and midpoint M_2 of A_2B_2 are respectively selected, and the angle θ between the line segment M_1M_2 and the vertical direction is the tower tilt angle.

4 Results

In the recent research of power tower, there is no data set for target detection of power tower by UAV aerial photography. Therefore, in order to solve the problem of ground target detection of power tower, this paper, on the one hand, adopts the Beidou navigation satellite-based UAV to assist shooting in several natural locations, and on the other hand, collects and sorts different types of power tower pictures from the network, a total of 625 pictures. At the same time, in order to enhance the robustness

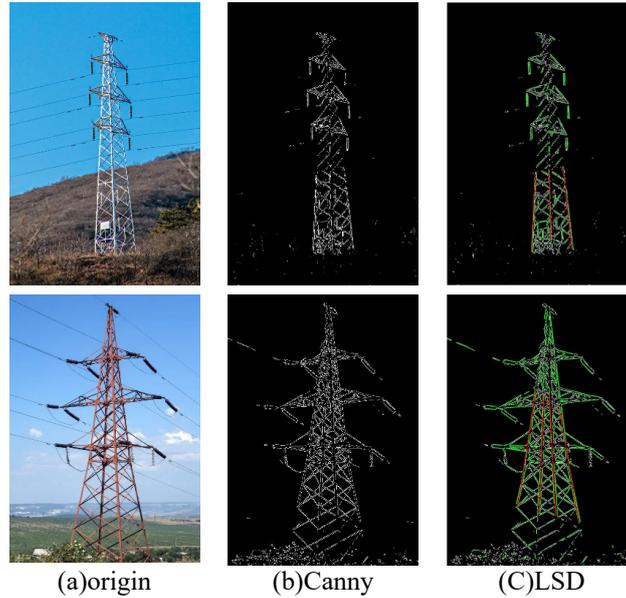


Figure 5 Test results of the tower

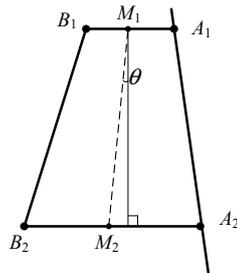


Figure 6 Schematic diagram of tower tilt angle

of the model, the image was scaled, translated, rotated, and other data enhancement operations. Finally, a total of 1500 images with 600×600 pixels were obtained, including 1000 images in the training set and 500 images in the test set.

The training and testing of this experiment were conducted on Ubuntu18.04 system with NVIDIA GeForce RTX 3090 graphics card, 24 gb video memory, CUDA version 10.2, and stochastic gradient descent method was used for training.

In order to compare the detection effects of several mainstream detection methods on power tower in UAV aerial images, this paper uses Faster R-CNN, SSD, YOLOv3 and the method in this paper to compare the target of power tower. The experimental results are shown in **Table 1**.

Table 1 Comparison results of power poles and towers

Model	Accuracy/%	Recall/%	F-measure/%
SSD	81.9	84.0	82.9
Faster R-CNN	80.2	84.6	82.3
YOLOv3	82.9	85.7	84.3
Ours	84.3	86.0	85.1

As can be seen from **Table 1**, the accuracy rate, recall rate and F value of the proposed method are 84.3%, 86.0% and 85.1%, respectively, which all reach the optimum in the compared model. Compared with SSD, F value is increased by 2.2%, compared with Faster R-CNN by 2.8%, and compared with YOLOv3 by 0.8%. It can be seen that the model proposed in this paper is effective.

Table 2 Tilt angle detection results

Angle range	0-5°	5°-25°	>25°
image number	260	195	45
accurate detection number	232	167	38
accuracy/%	89.2	91.8	84.4

In order to verify the overall accuracy of the power tower tilt monitoring method based on Beidou and improved YOLOv3 proposed in this paper, for all the picture data in the test set, the method of measuring the tower tilt angle one by one is used to compare with the detection angle of the tower recognition algorithm in this paper. If the difference between the measured Angle and the detection Angle is within 0.5°, the detection result is considered accurate. The test results of inclination are shown in **Table 2**. It can be seen from **Table 2** that the average accuracy of the model proposed in this paper reaches 87.4%, which can basically play an auxiliary role in monitoring the tilt of power tower rod.

5 Conclusion

In this paper, a tilt monitoring method of power tower based on BeiDou and improved YOLOv3 is proposed, and the tower identification algorithm and tilt detection algorithm are introduced in detail. For tower recognition, CFA and improved DarkNet53 were introduced on the basis of the original YOLOv3, and the F value reached 85.1%, which significantly improved the detection accuracy of small targets such as tower edges. For tilt detection, Canny edge detection algorithm and LSD straight line detection method were used to extract the center line and determine the tilt angle, and the final accuracy reached 87.4%. Based on the BeiDou navigation technology, the method proposed in this paper realizes the real-time and high-precision monitoring of the transmission line tower, which is of great significance to ensure the trouble-free operation of the tower.

6 References

- [1] Sun Q Q: Intelligent real-time monitoring technology and application of high voltage transmission tower based on BeiDou: 2022(01):58-63.

- [2] Zhao W, Wang Q, Shang K Y, et al: Electric Power Industry Precise Time-Space Service Network Based on BD Navigation System: Electronic production. *Electric Power Information and Communication Technology*, 2021, 19(07).
- [3] Girshick R: Fast r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1440-1448.
- [4] Ren S, He K, Girshick R, et al: Faster r-cnn: Towards real-time object detection with region proposal networks[J]. *Advances in neural information processing systems*, 2015, 28.
- [5] Redmon J, Divvala S, Girshick R, et al: You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.
- [6] Liu W, Anguelov D, Erhan D, et al: Ssd: Single shot multibox detector[C]//European conference on computer vision. Springer, Cham, 2016: 21-37.
- [7] Lin T Y, Dollár P, Girshick R, et al: Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.
- [8] Redmon J, Farhadi A: Yolov3: An incremental improvement[J]. *arXiv preprint arXiv:1804.02767*, 2018.
- [9] Liu S, Qi L, Qin H, et al: Path aggregation network for instance segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 8759-8768.
- [10] Hu L, Li Y: Micro-YOLO: Exploring Efficient Methods to Compress CNN based Object Detection Model[C]//ICAART (2). 2021: 151-158.
- [11] Lin T Y, Dollár P, Girshick R, et al: Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.
- [12] Shen X J, Du Y, Wang R, et al: Inclination measurement of transmission line tower based on terrestrial 3D lidar. *Journal of Electronic Measurement and Instrumentation*: 2017,31(04):516-521.
- [13] Lei X S, Sui Z H: Intelligent fault detection of high voltage line based on the faster R-CNN [J]: *Measurement*, 2019, 138: 379-385.
- [14] Guo J D, Chen B, Wang R S: Real-time inspection image of UAV power line tower based on YOLO: *Electric Power*, 2019,52(07):17-23.
- [15] Chollet F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1251-1258.