Journal of Computers Vol. 32 No. 2, 2021, pp. 124-136
doi:10.3966/199115992021043202011

Key Components Detection and Identification of Transmission Lines Based on an Improved CornerNet Network

Qing-Qing Zhang[1], Zhong-Jie Zhu[1*], Zhi-Feng Ge[2], Ming Gao[2], Yong-Qiang Bai[1], Ren-Wei Tu[1]

[1] Department of College of Information and Intelligence Engineering, Zhejiang Wanli University, Ningbo, China
731891448@qq.com, zhongjiezhu@hotmail.com, byq-163@163.com, 714450974@qq.com

[2] Department of Ninghai Power Supply Company Limited, State Grid Corporation of Zhejiang, Ningbo, China
37813913@qq.com,58782246@qq.com

Received 12 May 2020; Revised 22 August 2020; Accepted 15 September 2020

Abstract. Object detection for key components is an important means to ensure safe operation of transmission lines. However, the accuracy and robustness of object detection need to be further improved due to the influence of the weather and flight attitude of unmanned aerial vehicles. To solve these problems, this paper presents an object detection scheme based on an improved CornerNet deep network for transmission lines. Considering the characteristics of flight attitude, the background is divided into two types, and their own data sets are established for network training. Subsequently, the network structure is improved based on CornerNet deep networks to enhance feature propagation and improve network performance. In the test stage, the background classification is carried out first with a support vector machine, and then, object detection is implemented with the trained model for each background. The test results show the validity and feasibility of the scheme with an accuracy of 88.1% for the sky background and 92.3% for the ground background, which are 6.7% and 10.9% higher than those for the original network, respectively. In addition, through experiments simulating severe weather further verify the robustness of the proposed scheme.

Keywords: object detection, CornerNet, transmission lines

1 Introduction

Transmission lines, as an indispensable part of our power supply system, play a significant role in our production and life [1]. However, as transmission lines are exposed to the outdoors year-round, they are perhaps more likely to be destroyed, especially in a severe environment such as a thunderstorm. In addition, damage to key components (such as transmission towers and insulators) has greater potential to cause related residential-area power outages, discontinuing business for some enterprises and seriously affecting the normal production and life of residents [2]. Clearly, object detection for key components is an important means to ensure the safe operation of transmission lines [3]. Therefore, it is essential to carry out effective object detection for transmission lines to ensure safety of the power supply.

To ensure the security of transmission lines, it is crucial to solve the problems of object detection and identification for their key components. Object detection is an important research area in the image processing field and has a broad research foundation [4]. Presently, object detection methods can be divided into two categories: feature-based methods and deep-learning-based methods. Generally, for feature-based methods, the design and manual extraction of many artificial features are required. These methods have some defects, such as low efficiency and time consumption. With the development of

* Corresponding Author

technology, object detection has moved toward mainstream development of deep-learning-based methods. Alex et al. proposed Alexnet to establish the dominant position of neural networks in computer vision [5]. Simonyan proposed VGGNet, which showed that the network demonstrated good generalization performance and made it easier to engineer a project [6-7]. Ross et al. proposed Regions with Convolutional Neural Networks (R-CNN) features and Fast R-CNN to improve the speed and accuracy of object detection [8-11]. He et al. proposed the ResNet network, which solved the problem of vanishing gradient and promoted the development of object detection [12]. Joseph et al. proposed a series of You Only Look Once (YOLO) networks with a simpler network structure and a faster running speed [13-16]. These deep-learning-based methods avoided the limitations of artificial feature design and directly predicted the confidence of the bounding box and its categories [17-18]. Considering the particularity of the transmission lines, the object detection research in the transmission lines detection had also made the following attempts mainly on the traditional method, the hardware design, and the algorithm design of object detection, which will be roughly introduced as follows. First, the traditional method was manual detection, but there were some defects of high risk and low efficiency in field investigation. For hardware design, with the development of technology, the Unmanned Aerial Vehicle (UAV) had become an effective tool for detecting transmission lines. Yong et al. acquired the videos collected by UAV, which showed that more accurate, related information can be obtained through human monitoring [19]. Clearly, this semi-automatic approach was still time-wasting and labor-consuming. Rehak et al. used a robotic arm mount to simulate a UAV, building an artificial vision system to inspect transmission lines [20]. Finally, for the algorithm design of object detection, available methods mainly acquired images through UAV to recognize objects and obtain results to make judgments. For this goal, the design of a deep learning network was developing rapidly [21-22]. Zhang et al. proposed a new deep learning network RCNN-based object detection system for securing power transmission lines that would be suitable for detecting objects on power transmission lines [23]. This method showed that it significantly improved the recognition accuracy compared with the original Faster-RCNN. Nevertheless, the real-time performance of this method was not enough because large numbers of bounding boxes were needed and vast amounts of time were required. Han et al. proposed single-fault and multi-fault detection in aerial images, the backgrounds of which usually contained complex interference [24]. The shapes of the insulators were different due to the changes in filming angles and distance. The results of this method showed that it was more effective and efficient than the other insulators' fault detection methods. But this method had limitations in its application because it was only for insulator detection and not for all transmission lines detection. These methods were only suitable in a simple detection background with light workload. At present, there is still a lack of research on transmission line detection. Therefore, the requirement was urgent to realize efficient and automatic object detection for transmission lines. In terms of specific circumstances of transmission lines, object detection for their key components based on a deep learning network still had two major problems. On one hand, the flight attitude of UAV affected the backgrounds of images, which would affect the accuracy of object detection. Because of the constantly changing position of UAV, the backgrounds of transmission lines changed rapidly in captured video. These constantly changing backgrounds had an extraordinary effect on object detection for transmission lines [25]. On the other hand, the deep learning methods mentioned were almost based on anchor. These methods needed large numbers of anchor boxes to completely overlap with the ground truth boxes, which would lead to network structures becoming redundant [26-28]. In addition, this method created a giant imbalance between positive and negative anchor boxes and increased training time [29-31].

To solve these problems, this paper proposed a scheme of key components that included transmission lines based on an improved CornerNet network. First, the constantly changing backgrounds caused by UAV flight attitude increases the difficulty of detection. Therefore, Histogram of Oriented Gradient (HOG)+support vector machine (SVM) methods were adopted for image classification. In this paper, vast amounts of images were collected as a data set. Then, the images of the data set were manually annotated to establish a training data set (mainly composed of a transmission tower and insulators). Later, the training data set was classified by the HOG+SVM method and divided into a sky background subset and ground background subset. Meanwhile, we mainly improved the trunk network of CornerNet, which did not need the anchor boxes during the training and testing stage. Numerous experiments were conducted, inspiring our improved CornerNet network and accuracy of object detection.

Three main contributions of this paper can be summarized as follows: (1) Image classification can effectively avoid interference from background transformation. The image classification and object

detection were divided into two stages. The HOG+SVM method was applied to classify blended images, which can rapidly achieve background classification, so this will not impact the following object detection of transmission towers and insulators. (2) Optimization of network structure can effectively improve detection accuracy. In this paper, to achieve high precision, an optimized hourglass module was added. This hourglass module can effectively enhance the small object feature extraction. (3) The establishment of a data set contributed to the inspection of transmission lines as well as environmental noise simulation in severe weather. Since publicly known transmission line data sets were not available, a new data set of 5200 images of transmission towers and insulators was established in this paper. The experimental results verified the effectiveness of the proposed scheme. The background classification method in this paper was more conducive to extracting transmission lines in experiments and effectively eliminating the interference and influence of constant background changes from UAV flying attitude. It improved the detection accuracy of transmission lines' key components. Compared with the existing CornerNet method, the improved CornerNet network presented in this paper showed an accuracy of 88.1% and a detection accuracy of 6.7% increase for ground background; it showed an increased accuracy of 92.3% and 10.9% for sky background.

The remaining paper is organized as follows. Section 2 describes the proposed method, including the establishment of the data set, image classification, and optimization of the CornerNet network structure. Section 3 presents the experimental results and analysis. Finally, Section 4 presents conclusions.

2 Proposed Scheme

The proposed scheme consisted of two stages: training and testing. During the training stage, the data set was built by manual collection and UAV image capture. Preset image enhancement technology was adopted to improve the robustness of key component detection. The data set was then divided into a subset of sky and ground according to images of background information by an SVM classifier. Later, the two subsets were trained through the improved CornerNet network, and the background models were obtained, including the sky background model and the ground background model for object detection. During the testing stage, the test image was first classified by an SVM classifier according to their backgrounds. Following this was the model selection based on the results of the previous step; then the key components of transmission lines were detected in the test images by the sky background trained model or the ground background trained model. A diagram of the proposed scheme is shown in Fig. 1.

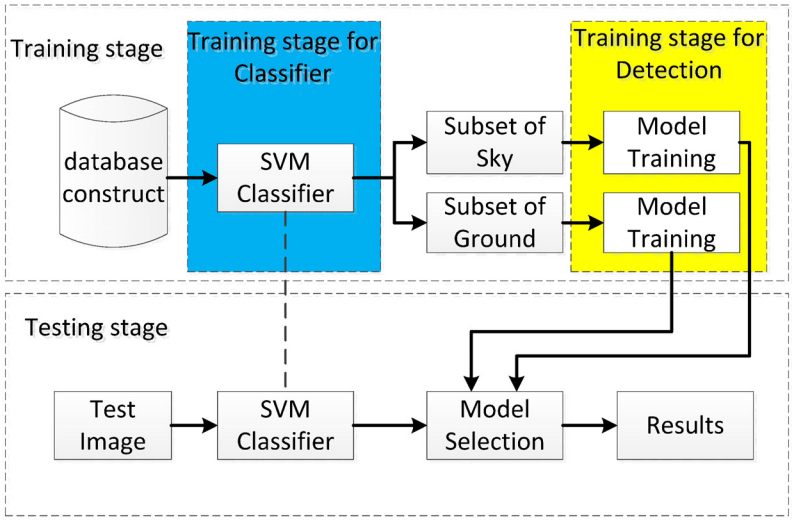


Fig. 1. Diagram of the proposed scheme

2.1 Establishment of Data Set

Generally, the detection effect of a deep neural network is limited by the size of the data set. At present, there are no data sets for transmission tower and insulators in public. To complete the key components

detection, images of transmission lines were collected by UAV and manually. A data set of 5200 images, with an average size of 4000 × 3000 pixels, was established to complete transmission-line detection. Then, 4000 images were randomly selected as the training data set, accounting for approximately 76% of the data set. The remaining 1200 images as the test set accounted for approximately 24% of the data set. Image augmentation technology was adopted for the data set to improve the robustness of the model. The technology produces a series of random changes to the original images to generate similar but different samples, thereby expanding the size of the data set and reducing the dependence of models on certain attributes. This response increased the generality of the model. In this paper, the methods of image augmentation technology include panning, zooming, horizontal flipping, and color transformation. Each augmented image was obtained by a lot of random combination transformations of the edge image. To improve the recognition ability of the model, this paper also added some unlabeled images (negative samples) in the data set. In addition, these images covered different lighting, shooting angles, resolutions and detection backgrounds, which meet the needs of sample diversity and pertinence. Some examples are shown in Fig. 2.

Fig. 2. Partial transmission towers and insulators

To improve the accuracy of object detection, transmission tower and insulators are manually labeled with care. During labeling, the boxes were exactly selected to enhance detection robustness. Sometimes, due to their shape similarity to towers, the tree area was easily mistakenly detected as a tower. As shown in Fig. 3, the yellow bounding box was the ground truth of the transmission towers, and the red bounding box was the ground truth of the insulators.

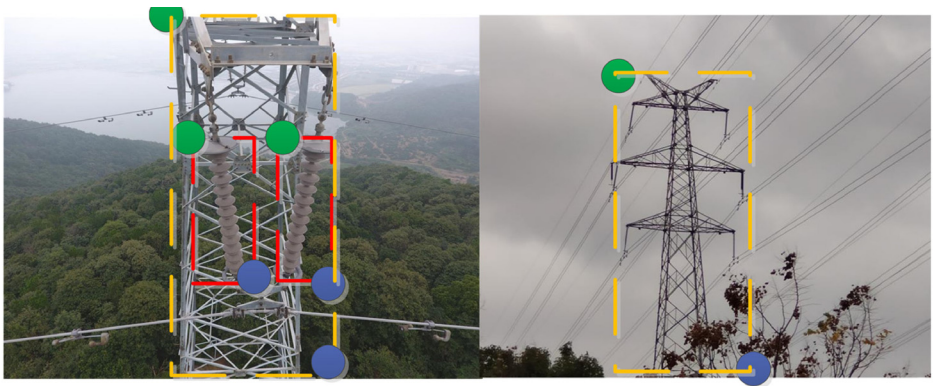


Fig. 3. Labeling of data set

2.2 Image Classification

We studied the reasons for the influence of complex backgrounds on the detection of transmission lines. For a more realistic complex background, common noises, such as Gaussian noise and salt-and-pepper noise, were employed to simulate an attack on the data set. We found that different noises, noise intensity, and composite noise had influence on the accuracy of object detection. Numerous studies of experiments showed that the higher the noise intensity and the more types of noise, the more destructive the detection accuracy would be. The Gaussian noise conforms to the Gaussian distribution as follows:

$$P(z)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(z-u)^2}{2\sigma^2}}. \tag{1}$$

Image classification was the innovation of the proposed scheme in this paper. When the data set was established in light of the UAV's flight attitude and human factors, the background of the data set was complex, which greatly affected the object detection effect. Therefore, the background of images was classified in this paper according to the image classification method and assigned the corresponding training model to detect. HOG was used to extract the preset sky and ground background features in the data set. The image information containing the preset ground background feature will be composed into the ground background subset just as it will with the sky background subset. An SVM classifier was used for image classification of the data set for three main reasons. First, through many experiments, it can be found that the complexity of the background interfered with the object detection and greatly affected the accuracy of detection. Second, image classification could preprocess the data of a complex background into different subsets, such as sky background and ground background. Third, an SVM classifier is a generalized linear classifier that classifies data sets according to supervised learning. Its advantages are fast speed and good classification effects.

During the training stage, image preprocessing was performed on the training data set. First, sky background and ground background were manually labeled. Later, we used HOG to extract sky and ground background features. Then, we used the sky and ground subsets for training to obtain the SVM classifier. During the testing stage, the test images would be judged by the trained SVM classifier to select the sky model or ground model for object detection in the next stage.

2.3 Optimization of Network Structure

In our scheme, a new simplified hourglass module was adopted to extract information for high precision. In the original CornerNet network, its backbone network consists of two Hourglass modules, and its depth is 104 layers [32]. This paper used a new simplified network consisting of three hourglass modules, with a depth of 54 layers. This hourglass had fewer parameters than the original one. Therefore, it responded faster than the one during the training stage. The time of object detection depended on the depth of the layers in the network. The deeper the layers, the longer the time required to train the model and the more calculation required [33]. An hourglass module was added, and the number of network layers of the hourglass module was optimized, changing its depth to 54 layers, as shown in blue in Fig. 5. The original CornerNet architecture is shown in Fig. 4.

The CornerNet network structure includes: the first convolution layer unit, the input unit, the residual unit, the first hourglass unit module, the convolution, the activation function unit, the network-in-network unit, the second residual unit, the second hourglass module, the third hourglass module, the first corner pooling layer unit, and the second corner pooling layer unit.

During the input-unit testing stage, the test image of the information transmission lines will link to the convolution layer unit. The residual until will in turn describe the first convolution layer cell size of the input image information. It will extract image information from the reduced first hourglass unit module, the convolution, the activation function unit, the network-in-network unit, the second residual unit, the second hourglass module, and the third hourglass module features. The third hourglass module outputs feature image information through the first corner pooling layer unit and the second corner pooling layer unit.

Further, the first corner pooling layer unit includes the first heatmaps, the first embeddings, and the first offsets. The second corner pooling layer unit includes the second heatmaps, the second embeddings, and the second offsets. The input and output information of each layer of the improved CornerNet is shown in Table 1.

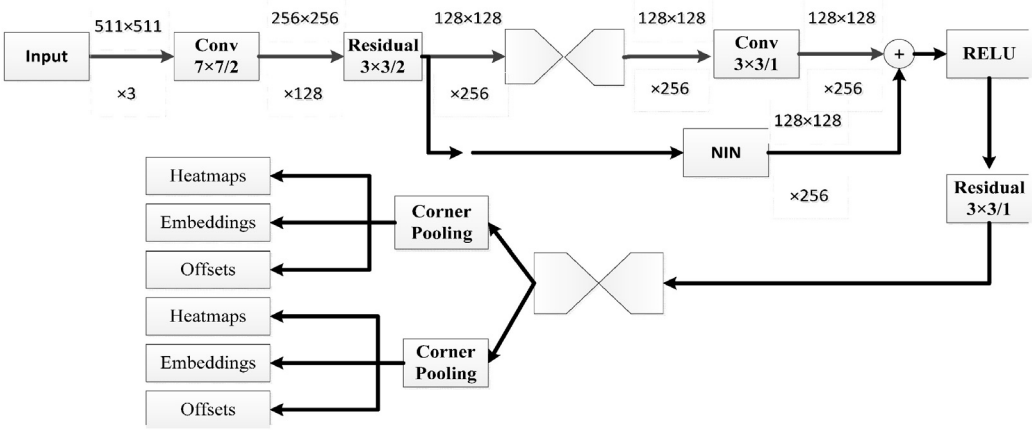


Fig. 4. Optimization of original CornerNet network structure

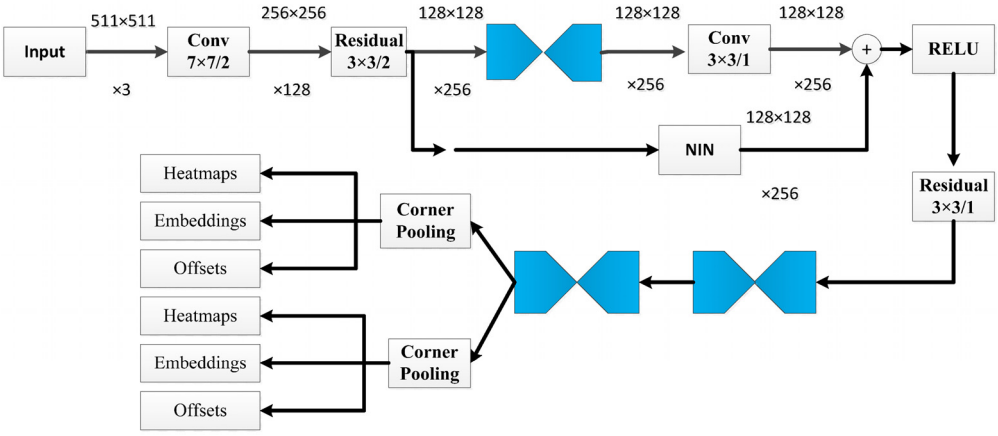


Fig. 5. Optimization of improved CornerNet network structure

Table 1. CornerNet network structure

Network layer name	Input	Output	Feature map size
Input	\	3	511*511
Conv7-BN-ReLU	3	128	256*256
Residual-s=2	128	256	128*128
Residual-s=2	256	256	64*64
Hg-up	256	256	64*64
Hg-max	256	256	64*64
Hg-low1	256	384	32*32
Hg-low2	384	384	16*16
Hg-low3	256	256	32*32
Hg-up2	256	256	64*64
Conv	256	256	64*64
Inter	256	256	64*64
Hg-up	256	256	64*64
Hg-max	256	256	64*64
Hg-low1	256	256	32*32
Hg-low2	384	384	16*16
Hg-low3	256	384	32*32
Hg-up2	256	256	64*64
Conv	256	256	64*64
Tl/Br_module	256	256	64*64
Tl/Br_heats	256	256	64*64
Tl/Br_tags	256	1	64*64
Tl/Br_offs	256	2	64*64

Two groups of heat maps were predicted, one for the top-left corner and the other for the bottom-right corner. Each group has two channels, which are respectively used to detect and identify transmission towers and insulators. For each corner, there is a positive ground-truth position, and the other positions are negative. During training, to avoid a false corner generated due to proximity to the ground-truth position, the false corner can also generate a frame position, overlapping with the ground truth. Therefore, to exclude those false corner-generated frames, CornerNet must set a certain intersection over union (IOU). In this paper, when detecting the transmission towers and insulators, IOU is set to be greater than 0.7. After the radius is determined, the attenuation from the center to the surrounding area is Gaussian formula, and the ground-truth position is in the middle, where σ = 1/3. Formula (2) is the loss function of the corner prediction.

$$F\det = \frac{-1}{n}\sum_{c=0}^{A}\sum_{x=1}^{B}\sum_{y=0}^{C}\begin{cases}(1-Pcxy)^{\alpha}\log(Pcxy) & if\ Ycxy=1\\ (1-Xcxy)\sigma^{\beta}(Pcxy)^{\alpha}\log(1-Pcxy) & if\ Ycxy\neq 1\end{cases}. \tag{2}$$

where the letter meaning is: *Pcxy* is the score of a position (x, y) of the class *A* transmission lines in the predicted heatmaps, *Ycxy* is ground truth. *N* is the number of objects in the image, and α and β are super parameters. *B* is the height of the image, and C is the width of the image.

The prediction module starts with a modified remnant, where CornerNet replaces the first convolution module with a corner pooling module. The modified residuals are followed by a convolution module. CornerNet has multiple branches for predicting heat maps, embeddings, and offsets.

During the training, to determine whether the pixel is in the top-left corner, you need to look horizontally to the right at the top edge of the object and vertically to the bottom at the left edge of the object. Therefore, the method of corner pooling and prior knowledge is adopted to better position the corner. If we want to make sure that (a, b) is in the top-left corner, let (a) be the input feature mapping of the top-left pooling layer, $Y_{h_{ab}}$ and $Y_{k_{ab}}$ should be the vectors of Y_h, and Y_k should be in the position of (a, b) respectively. The feature mapping of L×D, the corner pooling layer should be first of all. All the feature vectors Y_h between (a, b) and (a, L) in the maximum pooling should be the feature vectors h_{ab}, and then all the feature vectors Y_k between (a, b) and (D, b) in the maximum pooling should be the feature vectors k_{ab}. Finally, add h_{ab} and k_{ab}. This calculation can be expressed by formula (3) and (4):

$$h_{ab} = \begin{cases}\max(Y_{h_{ab}}, h_{(a+1)b}) & if\ a < L\\ Y_{h_{Lb}} & if\ a \geq L\end{cases}, \tag{3}$$

$$k_{ab} = \begin{cases}\max(Y_{k_{ab}}, k_{(a+1)b}) & if\ b < D\\ Y_{k_{aD}} & if\ b \geq D\end{cases}. \tag{4}$$

We use a maximum pooling operation as shown in Fig. 6. Specifically, horizontal scanning from right to left and vertical scanning from bottom to top are adopted, and then, the overlapping parts of the two are added to complete the feature mapping of maximum pooling.

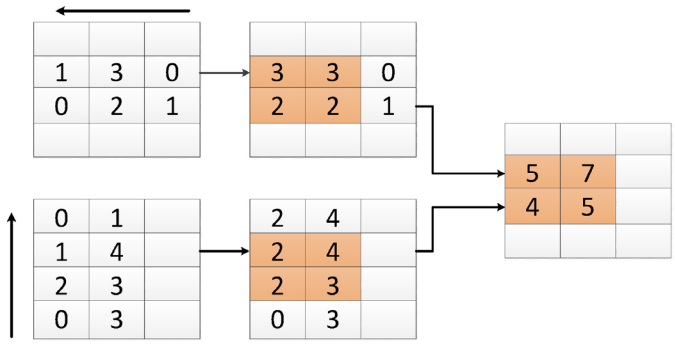


Fig. 6. Maximum pooling operation flow

3 Experimental Results and Analysis

In this paper, the experimental development environment is as follows: CPU: Intel i9-9920X, 3.5 GHz. GPU: NVIDIA GeForce RTX2080Ti 11G, Memory: 16G. Framework for deep Learning Networks: Hourglass-54. In addition, two types of detection objects are set, one is the transmission tower, the other is the insulator. The number of max iterations is 50000. A training model is generated every 5000 iterations, and a total of 10 models are generated after training. The resolution of the input image is 128×128. The whole training process adopts multiscale training.

Table 2. Set key training parameters

Scheme	Batch	Decay rate	Input	Learning rate	Max batches
CornerNet	4	10	128 × 128	0.00025	50000
Improved CornerNet	6	10	128 × 128	0.0001	50000

Through image classification, the data set is divided into sky background subset and ground background subset. The sky subset is shown in Fig. 7, and the ground subset is shown in Fig. 8.

Fig. 7. Example of a subset of the sky background

Fig. 8. Example of a subset of the ground background

Part of experimental results for key components of transmission lines are shown below. Fig. 9 is the partial detection effect picture under the sky background, and Fig. 10 is the partial detection effect picture under the ground background. These experiments show that our image classification combined with the improved CornerNet network is effective. The transmission tower and insulators in these figures are also accurately detected.

(a) (b)

Fig. 9. Example results of image detection for background of sky

(a) (b)

Fig. 10. Example results of image detection for background of ground

To be fair, the same test images are used for both the original CornerNet and improved CornerNet. The experimental results are shown in Fig. 11 and Fig. 12. Clearly, the original CornerNet will have high false detection and high rate of missed detection; that is, the insulators are missed detection, as shown in Fig. 11. After optimizing the network structure, the improved CornerNet networks are more precise in the detection of transmission towers and insulators.

(a) (b)

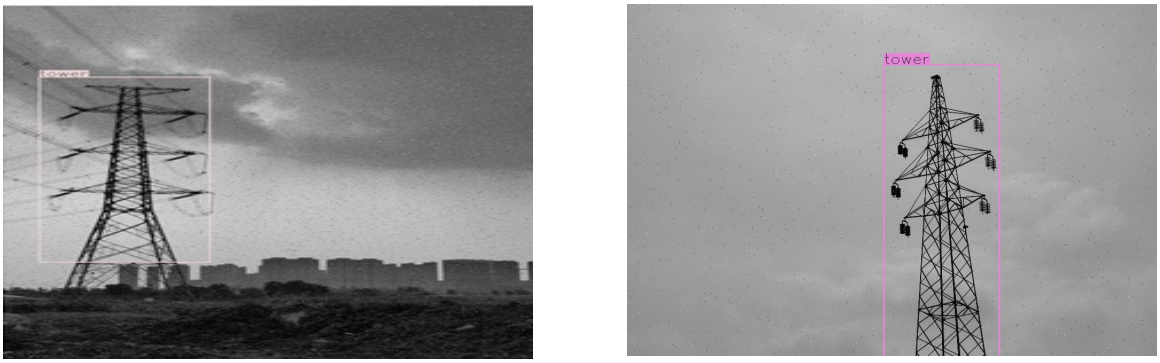
Fig. 11. Example results of the original CornerNet

(a) (b)

Fig. 12. Example results of the improved CornerNet

Gaussian noise is used to simulate severe weather conditions, as shown in Fig. 13(a). Gaussian noise is a type of noise in which probability density function is subject to Gaussian distribution. The common Gaussian noise includes undulating noise, cosmic noise, thermal noise, granular noise, and so on [34]. Salt-and-pepper noise (also known as impulse noise) is used to simulate severe weather conditions, as shown in Fig. 13(b), and it is the type of noise often seen in images. It appears as a random white or black dot, which may be black pixels in the bright region or white pixels in the dark region. The cause of the salt-and-pepper noise may be the sudden and strong interference of the image signal, the analog to digital converter, or the bit transmission error [35]. Through experiments simulating severe weather, the

improved CornerNet network is still effective in some bad weather.



(a) (b)

Fig. 13. Example results of the improved CornerNet under different types of noise interference

The overall comparison results are shown as follows. Both the accuracy and the recall rate of the improved CornerNet are effectively enhanced. Table 3 shows the results of the original CornerNet and the improved CornerNet on the test data set under the sky background. After the background image classification, the precision accuracy of the improved CornerNet is increased by approximately 6.7%, which substantially increases the applicability of the algorithm. The recall rate of the improved CornerNet is increased by 5.5%. Table 4 shows the results of the original CornerNet and the improved CornerNet on the test data set under the ground background. After the background image classification, the precision accuracy of the improved CornerNet is increased by approximately 10.9%, which substantially increases the applicability of the algorithm. The recall rate of the improved CornerNet is increased by 4.8%.

Table 3. Comparison of detection results before and after background image classification

Background	Recall	Precision	Test data set
Undivided background (original CornerNet)	0.753	81.4%	600
Background of sky (improved CornerNet)	0.808	88.1%	600

Table 4. Comparison of detection results before and after background image classification

Background	Recall	Precision	Test data set
Undivided background (original CornerNet)	0.753	81.4%	600
Background of ground (improved CornerNet)	0.801	92.3%	600

To further verify the validity of the proposed method, the same training set and test set are used to train and test with the YOLO V3 network. The test results of the YOLO V3 network are compared with those of the improved CornerNet network in this paper, and the results are shown in Table 5 and Table 6.

Table 5. Comparison of different network detection effects

different network	Recall	Precision	Test data set
YOLO V3	0.420	82.0%	600
Background of sky	0.808	88.1%	600
Background of ground	0.801	92.3%	600

Table 6. Comparison of different noise detection results

Noise types	Recall	Precision	Test data set
Gaussian noise in background of sky	0.890	74.2%	600
Pepper noise in background of sky	0.645	78.4%	600
Gaussian noise in background of ground	0.786	76.9%	600
Pepper noise in background of ground	0.645	74.3%	600

As shown in Table 5, the precision of YOLO V3 network is only 82.0%, and the recall of the YOLO V3 network is only 0.420. The improved CornerNet in this paper is better than YOLO V3 in precision and recall rate, clearly. Additionally, we used different types and sizes of noise to simulate the bad weather attack on the test data set. The effect of the detection is shown in Table 6. In experiments simulating bad weather, Gaussian noise and salt-and-pepper noise have a slight influence on the improved CornerNet, but our improved CornerNet still has the high-detection effect. Therefore, our proposed scheme has robustness to some degree.

The scheme of this paper has encouraged test results and is efficient and practicable. As an object detection attempt at precision and recall, the improved CornerNet in this paper has been encouraged. Our method is still efficient, despite the bad weather. In the face of the UAV's constantly changing posture in the background, the improved CornerNet still has high robustness.

4 Conclusions

The proposed scheme of key components detection and identification of transmission lines was based on the improved CornerNet deep network, which can eliminate the potential risk by detecting key components of transmission lines and efficiently issuing an early warning. For the sky background, the detection accuracy increased by 6.7%, and for the ground background, the detection accuracy increased by 10.9%. Later, the SVM classifier was used in the image classification, which can quickly perform the task of image classification. The scheme about the improved CornerNet was proposed, and it could improve the accuracy of transmission line object detection. Then, the data set about transmission towers and insulators was established. Finally, the trained model was used for simulation tests. The experimental results showed that the detection accuracy of the improved CornerNet was 88.1% for the sky background, and the accuracy was 92.3% for the ground background. It showed high accuracy and could meet the requirements of real-time performance. Following these results, the method showed good detection effects and strong robustness in the bad weather simulation.

The scheme of this paper had encouraged test results and was efficient and practicable as an object detection attempt at the precision and recall. However, many aspects, such as the image enhancement technique used in the data set, image data acquisition, and CornerNet network structure optimization, must be further studied. In future research, optimizing the network structure to reduce its computation and accomplish end-to-end network will be the main direction of our work.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (Grant No. 61671412), Zhejiang Provincial Natural Science Foundation of China (Grant No. LY19F010002, LY21F010014), Natural Science Foundation of Ningbo, China (Grant No. 2018A610053, 202003N4323), Ningbo Municipal Projects for Leading and Top Talents (Grant No. NBLJ201801006), General Scientific Research Project of Zhejiang Education Department (Grant No. Y201941122), Ningbo Municipal Project of Science and Technology Beneficial to People (Grant No.2017C50011), Innovation and Consulting Project from Ninghai Power Supply Company, State Grid Corporation of Zhejiang, China. School level Scientific Research and Innovation Team Project, and Fundamental Research Funds for Zhejiang Provincial Colleges and Universities.

References

[1] C. Yang, Improve efficiency and anti-offset using new pot ferrite core in wireless power transmission system, Electrical Engineering 101(3)(2019) 911-919.

[2] J. Boyuan, C. Argyropoulos, Nonreciprocal transmission in nonlinear PT-symmetric metamaterials using epsilon-near-zero media doped with defects, Advanced Optical Materials 23(2019) 1901-1915.

[3] W. Wu, Sequence-impedance-based stability comparison between VSGs and traditional grid-connected inverters, IEEE Transactions on Power Electronics 34(1)(2019) 46-52.

[4] L. Mohsen, Transaction protocol verification with labeled synchronization logic, Springer Cham 2019. doi: 10.1007/ 978-3-030-20652-9_19.

[5] A. Krizhevsky, I. Sutskever, G. Hinton, ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 60(6)(2017) 84-90.

[6] T. Chunwei, Attention-guided CNN for image denoising, Neural Networks 124(2020) 117-129.

[7] S. Junwon, L. Duque, J. Wacker, Drone-enabled bridge inspection methodology and application. Automation in Construction 94(2018) 112-126.

[8] S. Ren, Faster R-CNN: Towards real-time object detection with region proposal networks, IEEE Transactions on Pattern Analysis and Machine Intelligence 39(6)(2017) 1137-1149.

[9] Z. Cai, N. Vasconcelos, Cascade r-cnn: Delving into high quality object detection, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2018) 6154-6162.

[10] B. Cheng, Y. Wei, H. Shi, R. Feris, J. Xiong, T. Huang, Revisiting rcnn: On awakening the classification power of faster rcnn, In Proceedings of the European Conference on Computer Vision (ECCV) (2018) 453-468.

[11] W. Kelong, W. Zhou, Pedestrian and cyclist detection based on deep neural network Fast R-CNN, International Journal of Advanced Robotic Systems 16(2)(2019). doi: 10.1177/1729881419829651.

[12] K. He, Deep residual learning for image recognition, In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2016) 770-778.

[13] J. Redmon, S. Divvala, R. Girshick, You only look once: Unified real-time object detection, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016) 779-788.

[14] J. Redmon, A. Farhadi, YOLO9000: better, faster, stronger, In proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI. New York (2017) 6517-6525.

[15] S. Pang, A novel YOLOv3-arch model for identifying cholelithiasis and classifying gallstones on CT images, PLOS ONE14(6)(2019).

[16] T. Yunong, Apple detection during different growth stages in orchards using the improved YOLO-V3 model, Computers & Electronics in Agriculture 157(2019) 417-426.

[17] N. Bodla, B. Singh, R. Chellappa, L.-S. Davis, Soft-nms–improving object detection with one line of code, In Proceedings of the IEEE International Conference on Computer Vision (2017) 561-556.

[18] F. Chollet, Xception, Deep learning with depthwise separa-ble convolutions, In Proceedings of the IEEE conference on computer vision and pattern recognition (2017) 1251-1258.

[19] Z. Yong, R. Zhang, T.-J. Lim, Wireless, Communications with unmanned aerial vehicles: opportunities and challenges, IEEE Communications Magazine 54(5)(2016) 36-42.

[20] M. Rehak, J. Skaloud, Time synchronization of consumer cameras on Micro Aerial Vehicles, ISPRS Journal of Photogrammetry and Remote Sensing 123 (2017) 114-123.

[21] O.-A Menendez, M. Perez, F.A.A Cheein, Vision based inspection of transmission lines using unmanned aerial vehicles, In Proceedings of the 2016 IEEE International Conference on Multisensory Fusion and Integration for Intelligent Systems (2016) 412-417.

[22] J. Huaizu, Joint salient object detection and existence prediction, Frontiers of Computer Science: Selected Publications from Chinese Universities 13(4)(2018) 778-788.

[23] W. Zhang, RCNN-based foreign object detection for securing power transmission lines (RCNN4SPTL), Procedia Computer Science 147(2019) 331-337.

[24] H. Jiaming, A method of insulator faults detection in aerial images for high-voltage transmission lines inspection, Applied Sciences 9(10)(2019) 2009-2031.

[25] X. Zhou, J. Zhuo, P. Krhenbühl, Bottom-up object detection by grouping extreme and center points, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (2020) 850-859.

[26] L. Jianan, Attentive contexts for object detection, IEEE Transactions on Multimedia 19(5)(2016) 944-954.

[27] W. Jian, Enhanced object detection with deep convolutional neural networks for advanced driving assistance, IEEE Transactions on Intelligent Transportation Systems 21(4)(2020) 1572-1583.

[28] C.-B. Hao, S.-C. Huang, An advanced moving object detection algorithm for automatic traffic monitoring in real-world limited bandwidth networks, IEEE Transactions on Multimedia 16(3)(2014) 837-847.

[29] L. Hei, J. Deng, CornerNet: Detecting Objects as Paired Keypoints, International Journal of Computer Vision (2020) 1-15. doi 10.1007/978-3-030-01264-9_45.

[30] X. Ding, Local keypoint-based faster R-CNN, Applied Intelligence 50(6)(2020) 3007-3022.

[31] C. Antonio, Tuning hyperparameters of a SVM-based water demand forecasting system through parallel global optimization, Computers and Operations Research 106(2019) 202-209.

[32] J.-F. Zhang, H. Hu, G. Shen, Joint Stacked Hourglass Network and Salient Region Attention Refinement for Robust Face Alignment, ACM Transactions on Multimedia Computing Communications and Applications 16(1)(2020) 1-18.

[33] Newell, Alejandro, K. Yang, J. Deng, Stacked Hourglass Networks for Human Pose Estimation, 2016. doi: 10.1007/978-3-319-46484-8_29.

[34] M.-E. Helou, S. Susstrunk, Blind universal bayesian image denoising with gaussian noise level learning, IEEE Transactions on Image Processing 29(2020) 4885-4897.

[35] K. Vivek, A. Samadhiya, Image de-noising for salt and pepper noise by introducing new enhanced filter, 2019. doi:10. 15242/IIE.E1213577.