



Research Article

Analysis on the Impact of Data Augmentation on Target Recognition for UAV-Based Transmission Line Inspection

Chunhe Song ,^{1,2,3,4} Wenxiang Xu,^{1,2,3,4,5} Zhongfeng Wang,^{1,2,3,4} Shimao Yu,^{1,2,3,4} Peng Zeng ,^{1,2,3,4} and Zhaojie Ju ,^{4,6}

¹State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China

²Key Laboratory of Networked Control Systems, Chinese Academy of Sciences, Shenyang 110016, China

³Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110016, China

⁴Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China

⁵University of Chinese Academy of Sciences, Beijing 100049, China

⁶School of Computing, University of Portsmouth, Portsmouth, UK

Correspondence should be addressed to Peng Zeng; zp@sia.cn and Zhaojie Ju; juzhaojie@sia.cn

Received 21 July 2020; Revised 17 August 2020; Accepted 24 August 2020; Published 30 September 2020

Academic Editor: Guang Li

Copyright © 2020 Chunhe Song et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Target recognition is one of the core tasks of transmission line inspection based on Unmanned Aerial Vehicle (UAV), and at present plenty of deep learning-based methods have been developed for it. To enhance the generalization ability of the recognition models, a huge number of training samples are needed to cover most of all possible situations. However, due to the complexity of the environmental conditions and targets, and the limitations of images' collection and annotation, the samples usually are insufficient when training a deep learning model for target recognition, which is one of the main factors reducing the performance of the model. To overcome this issue, some data augmentation methods have been developed to generate additional samples for model training. Although these methods have been widely used, currently there is no quantitative study on the impact of the data augmentation methods on target recognition. In this paper, taking insulator strings as the target, the impact of a series of widely used data augmentation methods on the accuracy of target recognition is studied, including histogram equalization, Gaussian blur, random translation, scaling, cutout, and rotation. Extensive tests are carried out to verify the impact of the augmented samples in the training set, the test set, or the both. Experimental results show that data augmentation plays an important role in improving the accuracy of recognition models, in which the impacts of the data augmentation methods such as Gaussian blur, scaling, and rotation are significant.

1. Introduction

Transmission line inspection plays a very important role in ensuring the safety of the power system. In recent years, with the development of Unmanned Aerial Vehicle (UAV) technology, UAVs have been successfully applied in power transmission line inspection. In the traditional transmission line inspection method based on UAV, under the control of the operators, the UAVs fly along the specific inspection route and use the cameras to collect images or videos at certain locations. These images or videos will be transmitted back to the operators to obtain the actual flight paths, the

environment around the UAVs, and the targets to be identified. In this method, target recognition mainly depends on human experience, which limits the efficiency and accuracy of it [1].

With the progress of artificial intelligence technology, increasing attention has been paid to UAV autonomous inspection technology. As shown in Figure 1, by using GPS and intelligent image recognition technology, UAV can realize autonomous planning of inspection path, automatic obstacle avoidance, and intelligent recognition of suspicious targets. It is clear that the accuracy of image recognition directly determines the success of UAV autonomous



FIGURE 1: UAV-based autonomous transmission line inspection.

inspection. In recent years, the development of the deep learning technology has greatly improved the performance of image target recognition [2–4]. The performance of deep learning models highly rely on the number of valid training samples; unfortunately, lack of valid training samples is a common problem in deep learning model training. Due to the complexity of the environmental conditions and targets, as well as the limitations of images collection and annotation, the valid samples usually cannot cover most of all possible situations. As a result, the generalization ability of the recognition models is reduced, making them hardly usable in the practical scenes. To overcome this issue, some data augmentation methods have been developed to generate additional samples for deep learning models training [5–10]. Although these methods have been widely used, currently there is no quantitative study on the impacts of them on target recognition, which greatly limits the usage of these data augmentation methods.

In this paper, taking insulator strings as the target, the impact of a series of widely used data augmentation methods on the accuracy of target recognition is studied. The reason of choosing insulator strings as the target is that, in the transmission line, insulator strings are widely installed and have a variety of types, which are also fault prone elements. Once an insulator string fails, the transmission line cannot work normally and a large area of power failure is caused, which poses a great threat to the safe and stable operation of the power system. The recognition of insulator strings is a very important step to detect the fault of them, and the recognition accuracy of the insulator strings directly determines the result of insulator string fault detection. However, due to the complex background, the various shapes of insulator strings, uncertain shooting parameters of the cameras, and the recognition of insulator string are still a very challenging work. Although the recent advanced deep learning-based methods provide a promising approach to solve this issue, they suffer the insufficient training sample issue, similar with many other practical scenes. Therefore, data augmentation methods are usually adopted when training a deep learning-based model for insulator string recognition. Currently, the widely used data augmentation methods include histogram equalization [1, 6, 8], Gaussian blur [1, 7], translation [1, 5, 10], scaling [9, 10], and rotation [1, 9, 10]. Although these data augmentation methods have been widely used, there is still a lack of quantitative analysis of the impact of different

augmentation methods on recognition results, which seriously hinders the further improvement of data augmentation impact.

To overcome this issue, this paper studies the impact of some widely used data augmentation methods on the accuracy of target recognition, including histogram equalization, Gaussian blur, random translation, scaling, cutout, and rotation. Extensive experiments are carried out, and it is found that data augmentation plays an important role in improving the recognition performance of the model when the dataset is small. Meanwhile, Gaussian blur, scaling, and rotation have a great impact on the target recognition performance.

The rest of this paper is arranged as follows. Section 2 reviews the related works, and Section 3 gives the details of different data augmentation methods and examples. Section 4 presents extensive experiments and analyzes the impact of data augmentation methods on target recognition, and Section 5 is the conclusion.

2. Related Works

The safe and stable operation of the power system is of great significance to human life. Intelligent analysis of the power system has always been a research hotspot [11, 12]. As a key element of the power transmission line, insulator string recognition and defect detection is a research hotspot. As an important part of the transmission line, the state of insulator string directly determines the operation safety of the transmission line. Therefore, insulator string recognition and fault detection has always been a research hotspot. Insulator string recognition can be carried out using traditional image processing methods or the recent advanced deep learning-based methods. Because the insulator string in the image does not always follow certain directions, the recognition methods need to detect the insulator string along all possible direction, which is very time consuming. To overcome this issue, Zhao et al. [13] propose an insulator string recognition method based on orientation angle detection and binary shape prior knowledge. Zhao et al. [14] propose an insulator strings recognition method in infrared image based on binary robust invariant scalable keypoints (BRISK) and vector of locally aggregated descriptors (VLAD).

With the development of deep learning technology, many neural networks or convolutional neural networks (CNNs) methods have been proposed [15–20]. Zhao et al. [15] adopt the VGG16 structure and replace the last three fully connected layers with a VLAD pooling layer, and a SVM is trained for binary image classification. Sadykova et al. [9] adopt the You Only Look Once (YOLO) model for insulator string recognition. Chen et al. [18] use the You YOLO V3 algorithm for insulator strings recognition. Meanwhile, they improve the image quality using a super-resolution method based on Super-Resolution Convolutional Neural Network (SRCNN). In [20], Kang et al. use a faster R-CNN network for insulator string recognition, and a deep multitask neural network for insulator string defect detection. Miao et al. [5] use a single shot multibox detector

(SSD) for insulator string recognition. Jiang et al. [6] use SSD for insulator string defect detection from the entire image, the multi-insulator image, and the single-insulator image and then adopt ensemble learning to combine the results. In [7], Sampedro et al. first propose Up-Net, a fully convolutional network (FCN) architecture, for insulator string segmentation, and then design a Siamese convolutional neural network- (SCNN-) based method for insulator string defect detection. Ling et al. [8] and Li et al. [10] use a faster R-CNN network for insulator string recognition and a U-net for insulator string defect detection. Tao et al. [1] propose a CNN cascading architecture for insulator string recognition and defect detection.

Similar with many other practical recognition tasks based on deep learning technology, the performance of deep learning models of insulator string recognition highly rely on the number of valid training samples. However, due to the complexity of the environmental conditions and targets, as well as the limitations of image collection and annotation, the valid samples usually cannot cover most of all possible situations. To overcome this issue, some data augmentation methods have been developed to generate additional samples for deep learning model training. For example, Miao et al. [5] adopt horizontal and vertical flip; Jiang et al. [6] adopt horizontal flip and gamma correction; Tao et al. [1] adopt affine transformation, insulator and new background fusion, Gaussian blur, and brightness transformation; Ling et al. [8] adopt augmentation methods including random flip, crop, random saturation, brightness, and contrast perturbation. Sadykova et al. [9] adopt many augmentation methods such as Gaussian noise, Gaussian blur, average blur, Median blur, rotation, scaling, addition, and multiplication. Li et al. [10] adopt more augmentation methods such as mirroring, rotation, affine transformation, Gaussian white noise, brightness and color transformation, and other data augment operations. Although these image augmentation methods have been widely used, at present, there is lack of quantitative study on the impacts of them on target recognition.

3. Data Augmentation

3.1. Histogram Equalization. Histogram equalization is to stretch the image nonlinearly and redistribute the pixel value of the image so that the gray histogram of the original image changes from a certain gray range to a uniform distribution in the whole range. For a RGB image in this paper, first, the image is converted from RGB space to HSV space, and then histogram equalization is performed on the V channel in HSV space as

$$p_r(r_k) = \frac{n_k}{n}, \quad k = 0, 1, \dots, L - 1, \quad (1)$$

where n is the total number of pixels in the image, n_k is the number of pixels with gray level r_k , and L is the total number of possible gray levels in the image.

Meanwhile, the pixels of the gray level r_k in the image can be mapped to the corresponding pixels of the gray level s_k in the output image by

$$s_k = \sum_{j=0}^k p_r(r_j) = \sum_{j=0}^k \frac{n_j}{n}, \quad k = 0, 1, \dots, L - 1. \quad (2)$$

Figure 2 gives two examples of histogram equalization, in which Figures 2(a) and 2(c) are original images, and Figures 2(b) and 2(d) are corresponding images after histogram equalization. From Figure 2, it can be seen that, compared with the original images Figures 2(a) and 2(c), the images after histogram equalization Figures 2(b) and 2(d) have a better visual impact. However, whether this impact can benefit the performance of target recognition is not clear, which will be verified in Section 4.

3.2. Gaussian Blur. Gaussian blur is to transform the weight of each pixel in an image according to the Gaussian distribution function using the weighted average value with the surrounding pixels:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2/2\sigma^2)}, \quad (3)$$

where σ is the standard deviation, and x and y are the coordinates of pixels in the Gaussian blur kernel. Then a Gaussian blurred image I_b can be obtained by

$$I_b = I \oplus G(x, y), \quad (4)$$

where \oplus is the convolution operator.

The impact of Gaussian blur depends on the standard deviation σ and the size of the Gaussian blur kernel. In this paper, the standard deviation of Gaussian blur is fixed as 5, and the sizes of the Gaussian blur kernels are set to $3 * 3$, $7 * 7$, and $11 * 11$, respectively. Figure 3 gives an example of an image blurred by different Gaussian blur kernels, where Figure 3(a) is the original image and Figures 3(b)–3(d) are blurred images resulted from Gaussian blur with the kernels $3 * 3$, $7 * 7$, and $11 * 11$, respectively.

The Gaussian blurred image is to simulate the defocusing effect when the target is not in the focus position of the camera during exposure, which is a common phenomenon. Therefore, it may be an effective data augmentation method.

3.3. Random Translation. Random translation is to keep the size of the image unchanged and move the whole image up/down/left/right for a certain distance. In this paper, the translation distance is random, but ensures that the insulator string will not be moved out of the image. Figure 4 gives two examples of random translation, in which Figures 4(a) and 4(c) are original images, and Figures 4(b) and 4(d) are corresponding images after random translation. Note that the vacant area resulted from random translation are filled by zeroes.

3.4. Image Scaling. Image scaling is to resize the image while keeping the image aspect ratio of the length and width unchanged. This strategy is used to simulate the effect of different focal lengths of the camera on the shooting results.

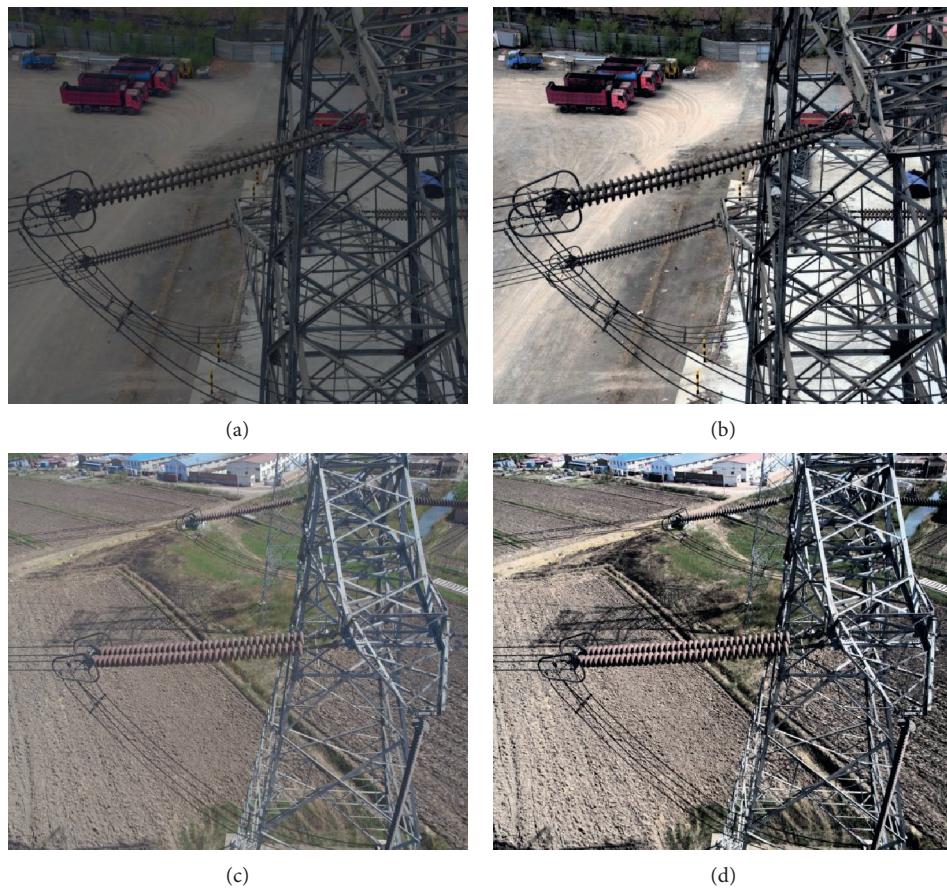


FIGURE 2: Images before and after histogram equalization.



FIGURE 3: Continued.



FIGURE 3: Images burried by different Gaussian blur kernels.

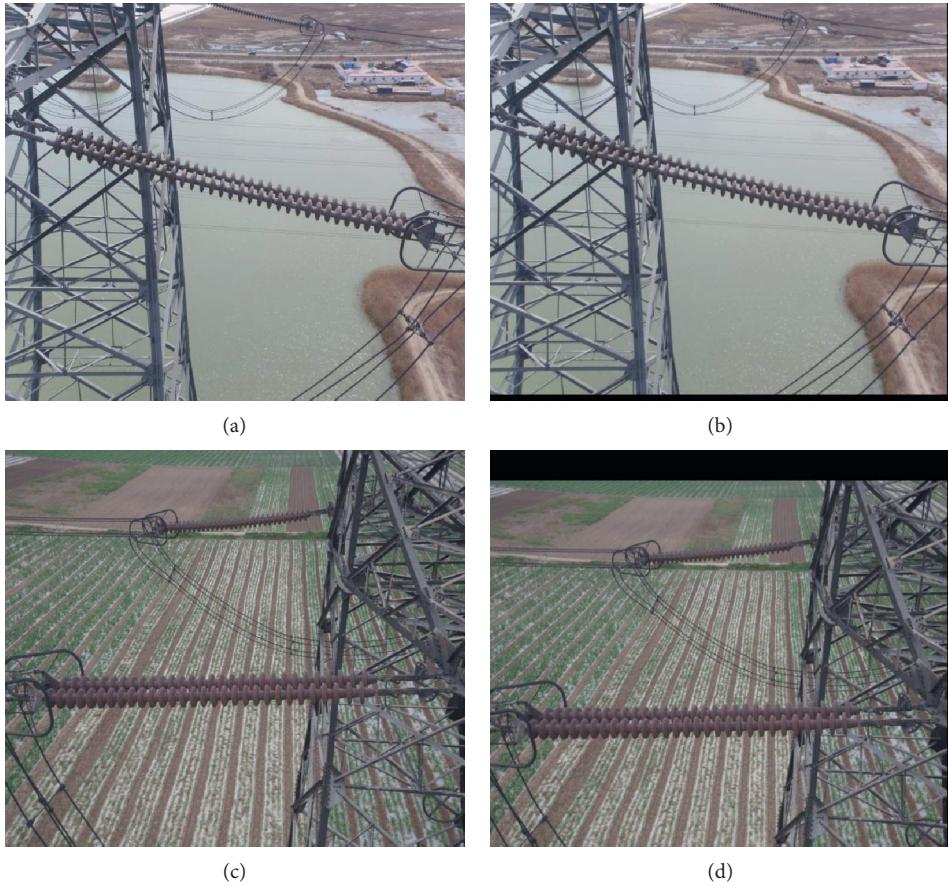


FIGURE 4: Images before and after random translation.

In this paper, two scaling ratios 0.5 and 2 are applied. Figure 5 gives two examples of image scaling, in which Figures 5(a) and 5(c) are original images and Figures 5(b) and 5(d) are corresponding images after image scaling. Similar with random translation, the vacant area resulted from image scaling are filled by zeroes. Note that, for some images, parts of insulator strings may be out of the images.

3.5. Image Cutout. Image cutout is the process of generating a new image by eliminating a certain region of the image. Cutout can simulate the situation that the target is partially occluded, which is a common phenomenon in a natural image. In this paper, image cutout is carried out by eliminating a region with the size of $100 * 100$. The region is randomly selected in any position of the image. Note that the

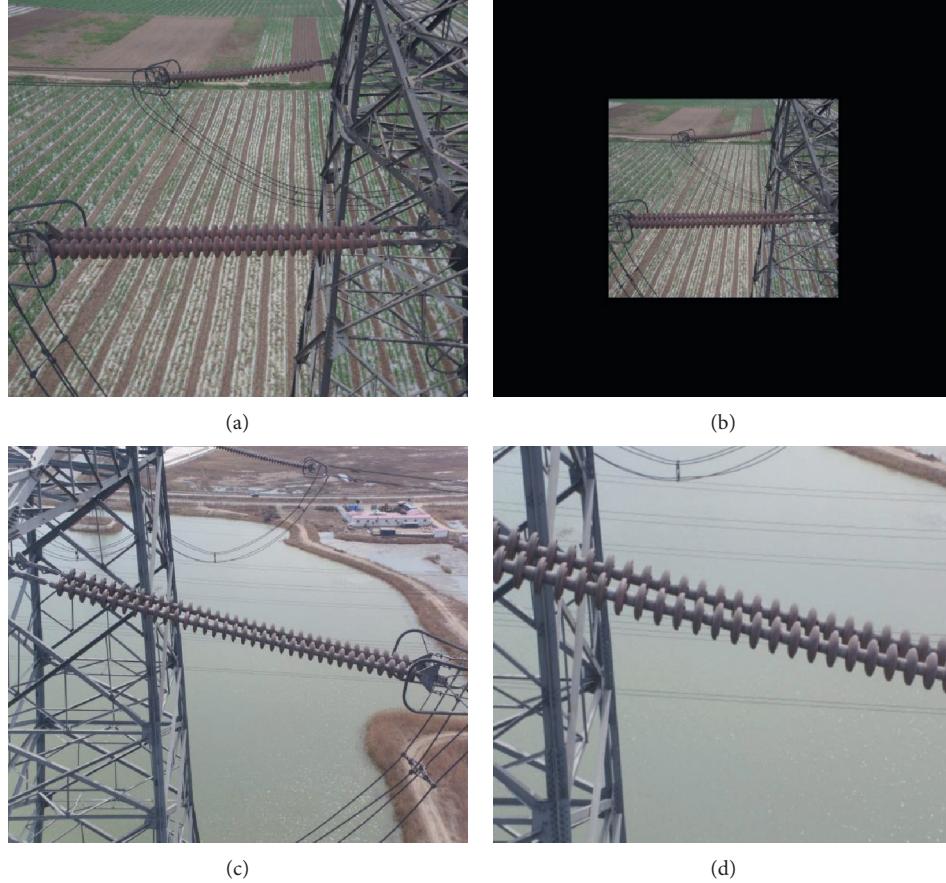


FIGURE 5: Images before and after scaling.

region may not be completely in the image, e.g., at the boundary of the image. Meanwhile, if less than 50% of the region appears in the image, then the region will be reselected. This strategy is used to ensure that the valid area of the eliminated region will not be significantly reduced at the boundary of the image. Figure 6 gives two examples of image cutout, in which Figures 6(a) and 6(c) are original images and Figures 6(b) and 6(d) are corresponding images after image cutout.

3.6. Image Rotation. Compared with general objects in natural images, the aspect ratio of insulator string is extremely large. The algorithms based on CNN are to search the region in the current image with high similarity with the labeled region in the training samples. When the algorithm based on CNN is applied to the identification of insulator strings, the overlapping area of insulator strings at different directions is greatly reduced due to the influence of the extremely large aspect ratio. Therefore, when the direction of the insulator string in the detected image is different from that in the training sample set, the detection rate will be greatly reduced. As a result, image rotation [1, 9, 10] has been adopted to increase the coverage of insulator string directions in the training set. In this paper, the relationship between the target rotation angle and recognition rate will be analyzed quantitatively. In order to achieve this goal, first,

the image is rotated to make the insulator string in the image in a horizontal position; second, the aspect ratio of the insulator strings in the images are calculated, and the images are divided into the large class and the small class according to the aspect ratio; finally, the images belonging to the large class and the small class are used for training and testing, respectively; for the test set image, it is rotated to -87° , -84° , ..., -3° , 0° , 3° , 6° , ..., 87° , and 90° , respectively, to test whether it can be recognized by the trained model.

Figure 7 gives examples of image rotation, in which Figure 7(a) is the original image and Figure 7(b) is a rotated image that the insulator string in the image is in horizontal position. Figures 7(c) and 7(d) are rotated images, where the insulator strings in the images are 3° and -3° from the horizontal.

4. Experiment and Discussion

The basic dataset used in this paper contains 848 insulator string images with a resolution of 1152×864 , including 600 images with insulator strings and 248 images without insulator string. The dataset is divided into the training set and test set with a ratio of 4:1, and the number of images of the training set and test set is 678 and 170, respectively.

The experiments are carried out on a server with ubuntu18.04, python 3.6, and rtx2080ti, and the deep learning framework is Caffe. In this paper, the software

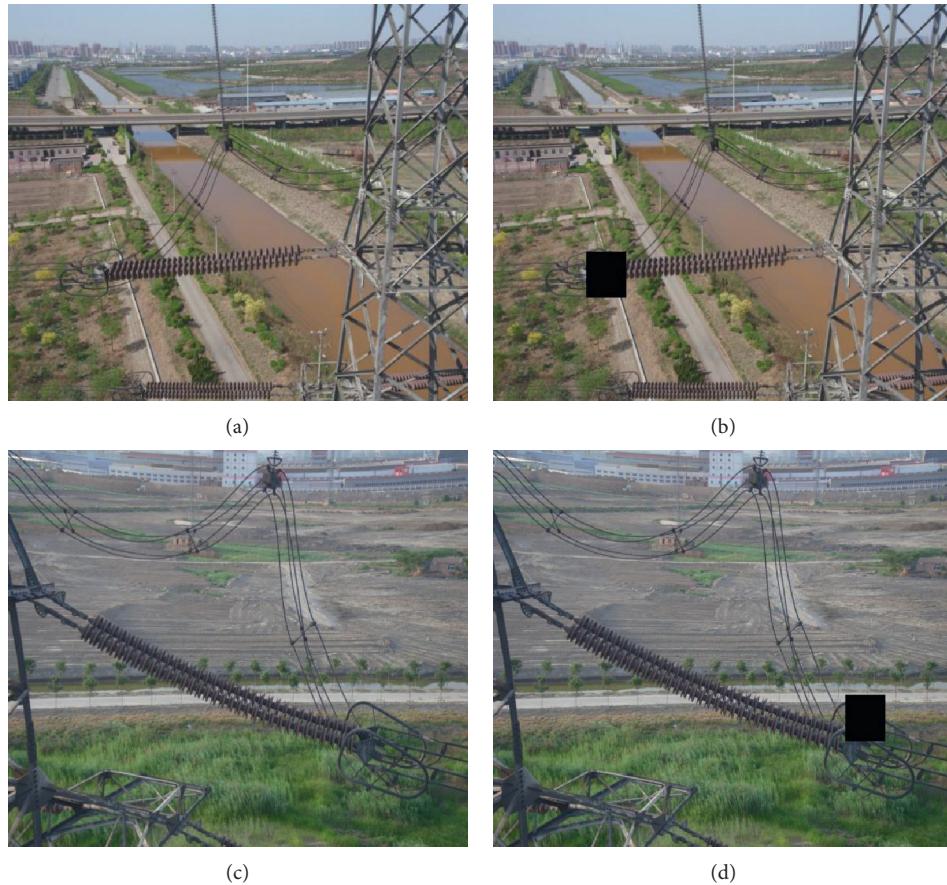


FIGURE 6: Images before and after cutout.



FIGURE 7: Continued.



FIGURE 7: Images before and after rotation.

LabelImg is used to label the 848 images in the basic dataset. The target recognition algorithm is faster RCNN [2], and the pretrained model is ZF. For the training process, the initial learning rate is set to 0.001, the weight attenuation coefficient is set to 0.0005, and the momentum value is 0.9.

4.1. Histogram Equalization. For this test, first the training set and test set are augmented by Gaussian blur with a kernel $3 * 3$, random translation, scaling ($\times 0.5$), and cutout, respectively. After that, the training set and test set are 4 times of the basic dataset, and these data are taken as the original dataset. In order to verify the impact of histogram equalization on recognition performance, the images in the original dataset are processed by histogram equalization, and then a new training set and test set are formed for training and testing. The test results before and after histogram equalization are shown in Table 1, where Ori, Aug, and All are the test results based on the original dataset, the augmented dataset, and all dataset, respectively.

It can be seen from Table 1 that when histogram equalization is used to augment the dataset, the recognition accuracy of the test set processed by histogram equalization is improved by 0.6%. For the original test data and augmented data, the recognition accuracy of the insulator string is improved by 0.22%, which shows that histogram equalization is a useful method to improve the accuracy of insulator strings recognition, although the impact is not significant.

4.2. Gaussian Blur. For this test, first the training set and test set are augmented by histogram equalization, random translation, scaling ($\times 0.5$), and cutout, respectively. After that, the training set and test set are 4 times of the basic dataset. In order to verify the impact of different Gaussian blur kernels on the recognition performance, the dataset is blurred by Gaussian blur with different kernels, and then a new training set and test set are formed. The test results with different Gaussian blur kernels are shown in Table 2, where K1, K2, and K3 are Gaussian blur with kernels $3 * 3$, $7 * 7$, and $11 * 11$, respectively.

TABLE 1: Test results before and after histogram equalization.

	Test set		
	Ori	Aug	All
<i>Training set</i>			
Ori	0.9002	0.8932	0.8975
All	0.9003	0.8992	0.8997

TABLE 2: Test results on Gaussian blur.

	Test set				
	Ori	K1	K2	K3	All
<i>Training set</i>					
Ori	0.9014	0.8924	0.7923	0.6038	0.8013
Ori + K1	0.9021	0.9004	0.8129	0.6307	0.8108
Ori + K2	0.9030	0.8999	0.8971	0.8100	0.8889
Ori + K3	0.9010	0.9006	0.8921	0.8922	0.8978
Ori + K1 + K2	0.9021	0.9012	0.8963	0.8136	0.8942
Ori + K1 + K3	0.9001	0.8994	0.8956	0.8798	0.8970
Ori + K2 + K3	0.9035	0.8995	0.8999	0.8821	0.8981
All	0.9033	0.9016	0.8975	0.8739	0.8985

It can be seen from Table 2 that the recognition accuracy of the model trained only with the original data on the test set images with $3 * 3$, $7 * 7$, and $11 * 11$ blur kernels decreases in turn; while for the model trained with blurred images and original images or only using the original image, when the test images are blurred with small Gaussian blur kernels (e.g., $3 * 3$), the difference of the recognition accuracy among different test datasets is not significant. It means that Gaussian blur with a small kernel has little impact on the result. However, when the training set contains images blurred by large kernels, the recognition accuracy on the test dataset is improved significantly. From the above analysis it can be seen that Gaussian blur is an effective method to improve the accuracy of insulator strings recognition, and the impact is significant.

4.3. Random Translation. For this test, first the training set and test set are augmented by histogram equalization, Gaussian blur with a kernel $3 * 3$, scaling ($\times 0.5$), and

cutout, respectively. After that, the training set and test set are 4 times of the basic dataset, and these data are taken as the original dataset. In order to verify the impact of random translation on recognition performance, the original dataset is randomly translated, and then a new training set and test set are formed for training and testing. The test results before and after random translation are shown in Table 3.

It can be seen from Table 3 that when the random translation method is used to augment the dataset, the recognition accuracy rate of the test set processed by random translation is improved by 0.36%, and the recognition accuracy rate of the insulator string is improved by 0.23% for the original test data and augmented data, which indicates that the recognition performance of the insulator string can be improved by using the random translation method.

4.4. Image Scaling. For this test, first the training set and test set are augmented by histogram equalization, Gaussian blur with a kernel $3 * 3$, random translation, and cutout, respectively. After that, the training set and test set are 4 times of the basic dataset, and these data are taken as the original dataset. To verify the impact of different scaling on recognition performance, the images in the original dataset are scaled by 0.5 and 2, respectively, and then a new training set and test set are formed for training and testing. The test results before and after scaling are shown in Table 4.

It can be seen from Table 4 that when using the model trained based on the original dataset, the recognition accuracy on the test images with the scale of 0.5 and 2 is obviously lower than on the original dataset. When the training set contains the images scaled by 0.5, the recognition accuracy of all dataset is improved. Furthermore, when the training set contains the images scaled by 2, the recognition accuracy of images scaled by 2 is improved. It shows that the recognition accuracy of images with different scaling ratios can be improved when the images with different scales are used for training.

4.5. Image Cutout. For this test, first the training set and test set are augmented by histogram equalization, Gaussian blur with a kernel $3 * 3$, random translation, and scaling ($\times 0.5$), respectively. After that, the training set and test set are 4 times of the basic dataset, and these data are taken as the original dataset. In order to verify the impact of image cutout on recognition performance, the images in the original dataset are processed by cutout, and then a new training set and test set are formed for training and testing. The test results before and after histogram equalization are shown in Table 5.

It can be seen from Table 5 that when using cutout to augment the original dataset, the recognition accuracy of the test set processed by cutout is improved by 0.56%. For the original test data and augmented data, the recognition accuracy of the insulator string is improved by 0.27%, which indicates that cutout is a useful method to improve the accuracy of insulator string recognition.

TABLE 3: Test results before and after random translation.

	Test set		
	Ori	Aug	All
<i>Training set</i>			
Ori	0.8996	0.8981	0.8981
All	0.9011	0.9017	0.9004

TABLE 4: Test results before and after scaling.

	Test set			
	Ori	$\times 0.5$	$\times 2$	All
<i>Training set</i>				
Ori	0.9026	0.7955	0.6230	0.7928
Ori + ($\times 0.5$)	0.9045	0.8121	0.6539	0.8016
Ori + ($\times 2$)	0.9024	0.7240	0.9080	0.8154
All	0.9030	0.8131	0.9087	0.9010

TABLE 5: Test results before and after cutout.

	Test set		
	Ori	Aug	All
<i>Training set</i>			
Ori	0.8997	0.8965	0.8996
All	0.9031	0.9021	0.9023

4.6. Image Rotation. For the basic dataset, first, rotate the image so that the insulator string in the image is in the horizontal position. Second, based on the statistics of the aspect ratio of insulator strings, it can be seen that the aspect ratio of insulator strings is from 1:1 to 14:1. In this paper, taking the aspect ratio 8:1 as the threshold, the images with the insulator string's aspect ratio greater than 8:1 are classified as class L, and those with the insulator string's aspect ratio greater than 8:1 are classified as class S. If there are both class L and class S insulator strings in an image, the image belongs to class L and class S at the same time. After that the images of class L and class s are divided into training set and test set according to the ratio of 4:1, and then histogram equalization, Gaussian blur, random translation, scaling, and cutout are used to expand the training set and test set of these two types of images, respectively.

During the test process, for an image, first, rotate it to make the insulator strings are at -87° , -84° , -3° , 0° , 3° , ..., 87° and 90° to the horizontal line, respectively; then, the recognition accuracies of the model for insulator strings in different angle are calculated and shown in Figure 8, where Figure 8(a) is the result of the L class images and Figure 8(b) is the result of the S class images. From Figure 8, it can be seen that the recognition accuracy of the model for insulator strings at 0° is the highest; however, with the angle deviation of 0° , the recognition accuracy of insulator string decreases dramatically. When the deviation angle is more than 12° , the model can hardly recognize the insulator string. Meanwhile, although the rotation angle is the same, the recognition

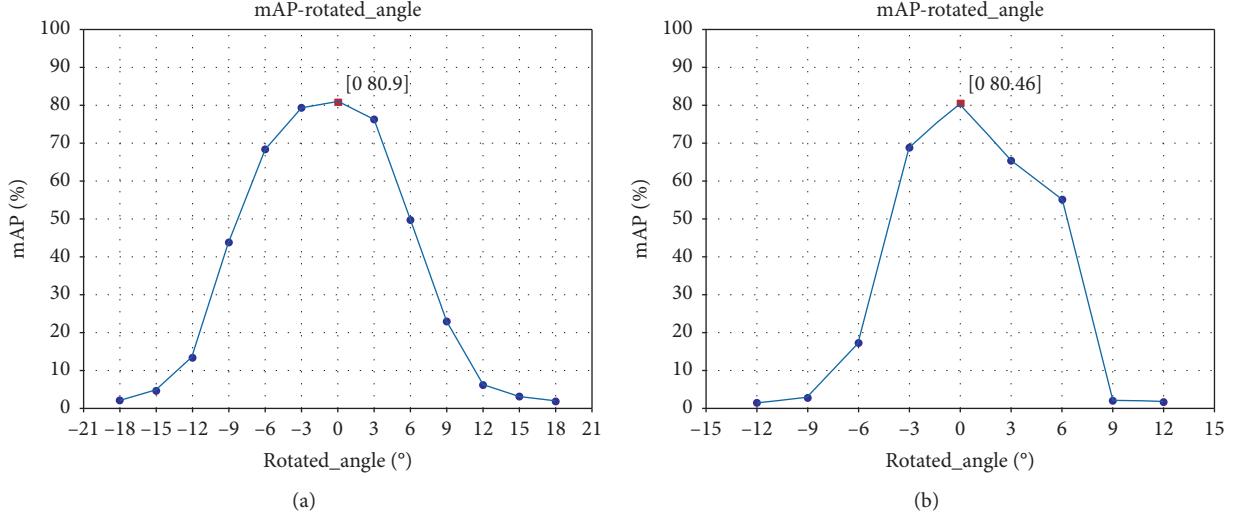


FIGURE 8: Recognition accuracy curve of two kinds of images.

accuracy is not the same due to different rotation directions. The results show that the general convolution neural network-based recognition method such as faster RCNN used in this paper cannot deal with the issue of target rotation well, and additional strategies are needed to solve the target rotation issue.

5. Conclusion

In this paper, the impact of data augmentation on target recognition for UAV-based transmission line inspection is analyzed. Based on extensive experiments on different data augmentation methods, it can be found that the data augmentation methods such as histogram equalization, random translation, and cutout can improve the target recognition accuracy, but the impact is not significant. Compared with above data augmentation methods, Gaussian blur, scaling, and rotation have a greater impact on the recognition performance of insulator strings. For rotation, the general convolution neural network-based recognition method such as faster RCNN used in this paper cannot deal with the issue of target rotation well, and additional methods should be adopted to solve this issue.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was partially supported by the National Key R&D Program of China under Grant 2018YFB1700200, National Nature Science Foundation of China under Grant

U1908212, State Grid Corporation Science and Technology Project (SG2NK00DWJS1800123), and State Grid Shanghai Electric Power Corporation's 2020 technology project "Research on Automatic Networking and Data Fusion Technology of Power IoT Sensors" (52090F1900BK).

References

- [1] X. Tao, D. Zhang, Z. Wang, X. Liu, H. Zhang, and D. Xu, “Detection of power line insulator defects using aerial images analyzed with convolutional neural networks,” in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 4, pp. 1486–1498, 2020.
- [2] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: towards real time object detection with region proposal networks,” in *International Conference on Neural Information Processing Systems*, vol. 50, pp. 91–99, 2015.
- [3] W. Liu, D. Anguelov, D. Erhan, and C. Szegedy, “SSD:Single shot multibox detector,” *ECCV*, vol. 50, pp. 21–37, 2015.
- [4] J. Redmon, S. Divvala, R. Girshick et al., “You only look once: unified, real-time object detection,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 50, pp. 779–788, 2016.
- [5] X. Miao, X. Liu, J. Chen, S. Zhuang, J. Fan, and H. Jiang, “Insulator detection in aerial images for transmission line inspection using single shot multibox detector,” in *IEEE Access*, vol. 7, pp. 9945–9956, 2019.
- [6] H. Jiang, X. Qiu, J. Chen, X. Liu, X. Miao, and S. Zhuang, “Insulator fault detection in aerial images based on ensemble learning with multi-level perception,” in *IEEE Access*, vol. 7, pp. 61797–61810, 2019.
- [7] C. Sampedro, J. Rodriguez-Vazquez, A. Rodriguez-Ramos, A. Carrio, and P. Campoy, “Deep learning-based system for automatic recognition and diagnosis of electrical insulator strings,” *IEEE Access*, vol. 7, pp. 101283–101308, 2019.
- [8] Z. Ling, “An accurate and real-time method of self-blast glass insulator location based on faster R-CNN and U-net with aerial images,” in *CSEE Journal of Power and Energy Systems*, vol. 5, no. 4, pp. 474–482, 2019.
- [9] D. Sadykova, D. Pernebayeva, M. Bagheri, and A. James, “IN-YOLO: real-time detection of outdoor high voltage insulators

- using UAV imaging,” in *IEEE Transactions on Power Delivery*, vol. 35, no. 3, pp. 1599–1601, 2020.
- [10] X. Li, H. Su, and G. Liu, “Insulator defect recognition based on global detection and local segmentation,” in *IEEE Access*, vol. 8, pp. 59934–59946, 2020.
 - [11] C. Song, W. Jing, P. Zeng, H. Yu, and C. Rosenberg, “Energy consumption analysis of residential swimming pools for peak load shaving,” in *Applied Energy*, vol. 220, pp. 176–191, 2018.
 - [12] C. Song, W. Jing, P. Zeng, and C. Rosenberg, “An analysis on the energy consumption of circulating pumps of residential swimming pools for peak load management,” in *Applied Energy*, vol. 195, pp. 1–12, 2017.
 - [13] Z. Zhao, N. Liu, and L. Wang, “Localization of multiple insulators by orientation angle detection and binary shape prior knowledge,” in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 22, no. 6, pp. 3421–3428, 2015.
 - [14] Z. Zhao, G. Xu, and Y. Qi, “Representation of binary feature pooling for detection of insulator strings in infrared images,” in *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 5, pp. 2858–2866, 2016.
 - [15] C. Song, P. Zeng, Z. Wang, T. Li, L. Qiao, and L. Shen, “Image forgery detection based on motion blur estimated using convolutional neural network,” *IEEE Sensors Journal*, vol. 19, no. 23, pp. 11601–11611, 2019.
 - [16] C. Song, P. Zeng, Z. Wang, H. Zhao, and H. Yu, “Wearable continuous body temperature measurement using multiple artificial neural networks,” in *IEEE Transactions on Industrial Informatics*, vol. 14, no. 10, pp. 4395–4406, 2018.
 - [17] Z. Zhao, X. Fan, G. Xu, L. Zhang, Y. Qi, and K. Zhang, “Aggregating deep convolutional feature maps for insulator detection in infrared images,” in *IEEE Access*, vol. 5, pp. 21831–21839, 2017.
 - [18] X. Gong, Q. Yao, M. Wang, and Y. Lin, “A deep learning approach for oriented electrical equipment detection in thermal images,” in *IEEE Access*, vol. 6, pp. 41590–41597, 2018.
 - [19] H. Chen, Z. He, B. Shi, and T. Zhong, “Research on recognition method of electrical components based on YOLO V3,” *IEEE Access*, vol. 7, pp. 157818–157829, 2019.
 - [20] G. Kang, S. Gao, L. Yu, and D. Zhang, “Deep architecture for high-speed railway insulator surface defect detection: denoising autoencoder with multitask learning,” in *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 8, pp. 2679–2690, 2019.