Journal of Physics: Conference Series

PAPER • OPEN ACCESS

SNIPER Based Multi-Target and Multi-Scale Aerial Image Processing Method

To cite this article: Yaocheng Li *et al* 2020 *J. Phys.: Conf. Ser.* **1659** 012003

SNIPER Based Multi-Target and Multi-Scale Aerial Image Processing Method

Yaocheng Li[1*], Weidong Zhang[2], Yingming Cai[2], Zhe Li[1] and Xiuchen Jiang[1]

[1] Department of Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[2] Weihai Power Supply Company of State Grid Shandong Electric Power Company

[*]Corresponding author's e-mail address: yaochengli539@gmail.com

Abstract. The maintenance of power transmission system appeals automatic inspection system to replace manual inspection. Inspection based on Unmanned Aerial Vehicle (UAV) generates a huge number of photos to be processed. These aerial photos have several challenging features: high density, small, scale variation. In order to ameliorate the performance on detection of important objects as insulators and pin bolts, an efficient multi-scale training algorithm called SNIPER is introduced. SNIPER enhanced Faster RCNN was trained on a powerline dataset for object detection. SNIPER provides a good average precision and average recall on large objects like insulators and medium objects with adequate annotated instances like pin bolts and dampers. However, SNIPER fails to locate missing pin or displaced pin, possibly due to their similarity to pin bolts. Future development of a SNIPER-based cascaded detection scheme could help detect defected small objects.

1. Introduction

Power transmission system plays a crucial role in modern society. While most part of it is in outdoor environment and thus suffers from deterioration, which may destabilize the power system and even cause power blackout[1].

Most vulnerable components of electrical pylons include, insulators, grading rings, pin bolts and dampers. Defects from insulators and grading rings can cause electric discharge and eventually electric failure, while loosened pin bolts and damaged dampers may cause mechanical disintegration of powerline from pylon.

Nowadays, power line inspection conducted by Unmanned Aerial Vehicle (UAV) generates an excessive quantity of photos. It is worth noting that these aerial photos of electrical pylons have several challenging features, such as:

- High density. With flexible mobility of UAV, numerous fasteners and insulators can be captured closely in a photo
- Small object. An insulator takes averagely 2% of total area, for pin bolts, 0.1% to 0.05%
- Scale variation. Several pylons can be captured in a same photo brought by a wide view of UAV. Scales of objects in a picture vary along with various distances.

Inspection task is very time consuming and tiring. Manual inspection will not be able to satisfy the demand of rapidly growing power transmission network[2]. The development of an automatic inspection system has become a vital necessity in this time.

Scientific literature has offered a wide scope on the development of an automatic inspection system to meet demands of power system and utilities. From handcrafted features to neural network, research on vision-based insulator inspection has been constantly conducted during the last years. In article [3], general Hough transform and RANSAC are employed to recognize and localize insulators from power station photos. Article [4]proposes a localization method of insulator based on Orientation Angle Detection and Binary Shape Prior Knowledge, its robustness can be questionable when met with different types of insulator or occlusion. Literature has widely used detectors based on convolutional neural network (CNN) to perform detection tasks, but is particularly difficult for CNN to detect small objects like pins and various defects because of their small nature[5]. In reality, images with defected equipment are rare, causing imbalanced distribution between normal and defected objects, which may lead to deficiency of anomaly detection algorithm. RetinaNet is utilized in [6]to tackle the aforementioned problem in the classification of normal and defected pin bolts.

Dampers attenuate the vibration of transmission line, so as to avoid mechanical failure as loosening fasteners. Usual dumper faults can be dislocation, sliding, missing, incompleteness etc. A comparative study of different backbones on Single Shot Multibox Detector (SSD) based on UAV is conducted on dampers, insulators and suspension clamps in[7].

In this work, an efficient multi-scale training algorithm is introduced to ameliorate is introduced to ameliorate the performance on multi-scale detection task, which is called SNIPER[8]. A multi-scale aerial image dataset is created. Important components of power transmission system are annotated in the dataset. SNIPER algorithm enhanced Faster RCNN detector is trained on this dataset. This article is organized as follows. In 2, the multi-scale feature of real-world data and SNIPER algorithm is presented. In 3, the dataset and the training process are described, and the trained model is evaluated. In 4, conclusions and perspectives are given.

2. SNIPER Algorithm

Figure 1. An image fromthe dataset (bounding boxes of insulators with green points at corners

As shown in Figure 1, several scales of insulators are present simultaneously in an image, evidently the same for all other fasteners, dampers etc. A detection algorithm of robustness to scales is necessary. Besides, backgrounds such as sky, urban constructions, mountains, rural scenes and greenery in power transmission images take up most pixels of an image. Multi-scale methods like image pyramid slows down the training because all pixels in the original image are processed multiple times [8] In multi-scale training, Fast-RCNN [9] upsamples and downsamples each proposal (regardless of its size) in the image. To upsample large objects and to downsample small objects are intuitively illogical concerning improvement of multi-scale detection performance. In

RCNN, objects are equally resized to a fixed 224*224 pixel size[10]. Thus, large objects are not upsampled and small objects are not downsampled. In this way, small objects do not lose information while being scaled and large objects shall not contain interpolated pixels which bring few information, both are brought to an appropriate scale to be processed by RPN. In the meanwhile, proposals of RCNN do not contain convolutional contextual information, unlike Fast-RCNN. It is enviable that less important background pixels are excluded while contextual information around objects remains. To that end, Scale Normalization for Image Pyramids with Efficient Resampling (SNIPER) was designed.

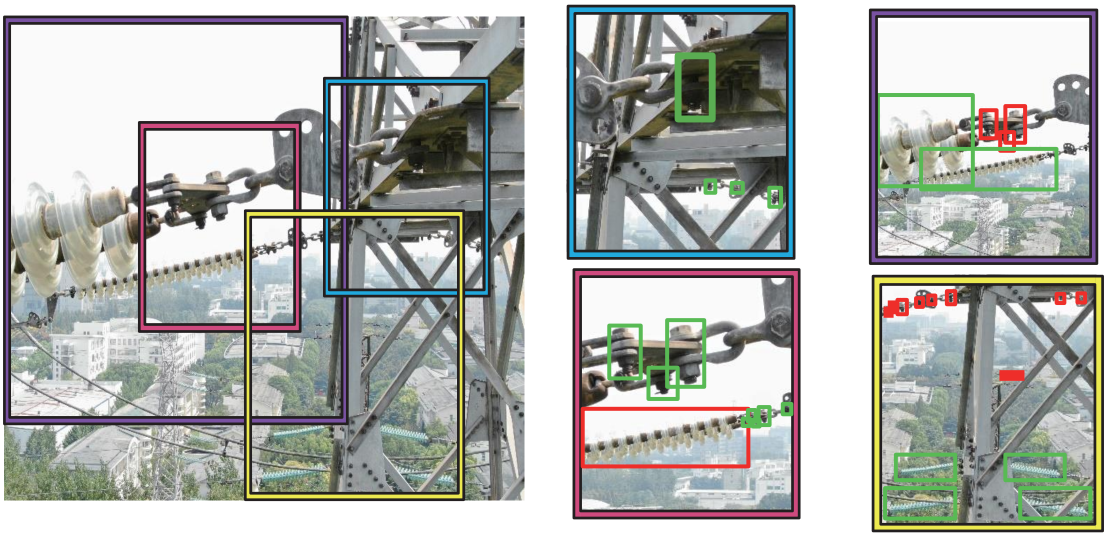

Figure 2. Illustration of positive chips selection process

SNIPER is a training algorithm which crops chips from multiple scales of an image. Cropped chips with scale-appropriate labels are sent to RPN. Positive chips contain scale-appropriate ground-truths with context, and ground-truths that are inappropriate are omitted. This cropping procedure is made to cover maximum amount of ground-truths. The generation of positive chips is illustrated in Figure 2: contextual regions (chips) are sampled adaptively in function of the appearance of objects inside the image. Left: The image, four cropped chips in original scale. Right: Down/up-sampling is processed regarding the size of the objects. Valid objects are shown in green and invalid objects in red rectangles will be rejected[8]. Negative chips are created to avoid false positives in background. Chips are then classified into negative chips and positive chips by a lightweight Regional proposal network (RPN), where positive chips are likely to contain objects while negative chips are simple backgrounds. Faster-RCNN detector then performs detection and classification on each positive chip.

3. Experiments

SNIPER was evaluated on our aerial photo dataset for object detection. The computer is equipped with Intel i9 9900K and GeForce RTX2080Ti with 11GB of memory. Backbone of the detector is ResNet101 pretrained on ImageNet. The batch-size is 12 and the learning rate is 0.0015. 80% of dataset is choosen randomly as training set, the rest is validation set. Loss function is cross-entropy.

3.1. Dataset

The dataset is composed of 604 aerial captured power transmission images with more than 7,000 bounding box labels in total. Insulators, pin bolts, dogbone dampers and grading ring have been annotated along with pin defects. Classes of dataset are presented in Figure 3. Classes labelled in the dataset.

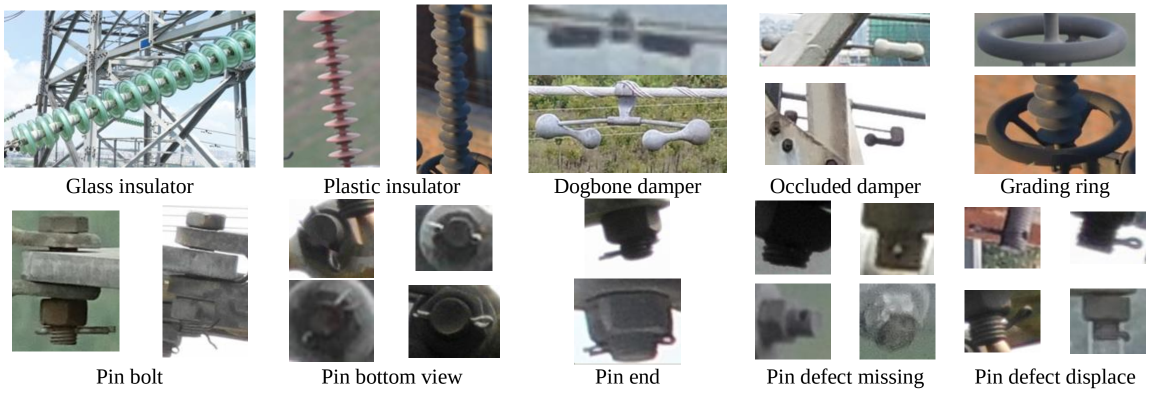

Figure 3. Classes labelled in the dataset

As for diagnostic of pin bolts, it can be difficult even for human to deduce whether a pin is properly installed from certain peculiar viewpoints, also it is influenced by light condition. Pin bolts are separated into 3 categories by different views, since these views of pin bolts are not similar to each other. Two pin defects, pin missing and pin displace, are included as well.

3.2. Result and discussion

Table 1. Statistics on labels of dataset

Label name	**mAP @0.5**[a]	**mAR @0.5**[a]	**quantity**	**Square root average area (pixel)**	**Average object scale/image scale**		
					width	***height***	***area***
Plastic insulator	0.6135	0.6747	394	413	6.807%	26.040%	2.054%
Glass insulator	0.6235	0.6451	2357	320	10.908 %	10.235%	2.045%
Grading ring	0.4710	0.5536	221	144	4.947%	3.632%	0.455%
Dogbone damper	0.5850	0.6306	1354	99	3.396%	2.300%	0.169%
Occluded damper	0.0351	0.2353	175	96	2.980%	2.416%	0.126%
Pin bolt	0.5873	0.6183	1845	88	2.046%	3.917%	0.117%
Pin bottom view	0.2485	0.2745	841	53	1.493%	2.350%	0.051%
Pin end	0.1119	0.0123	312	64	1.676%	2.425%	0.051%
Pin defect missing	0.0098	0.0351	238	40	1.026%	1.403%	0.019%
Pin defect displace	0.0040	0.1667	73	45	1.328%	1.631%	0.025%

[a] Mean average precision or recall at intersection over union at 50%

Table 1. shows that in the context of UAV photos of power line, the most interesting objects are generally small and dense. Glass and plastic insulators cover an average area of 2%, which is the largest in the dataset. Pin defects, which are among the most valuable objects, merely occupy about 0.02% of an image's area.

The trained detector performs best on the detection of large objects like insulators and medium objects of adequate annotated instances like pin bolts, dogbone dampers. Detector fails at small objects such as pin missing and occluded damper, which are de facto subclasses of pin bolts or dogbone. These classes are imbalanced in quantities, so as their contributions to the learning process can be too small compared to large objects. Besides, the similarity between pin defects and normal pins could be factors behind the difficulty to distinguish one another, the same as normal dampers and occluded dampers.

Figure 4. Detection result on an aerial image with insulators of different scales

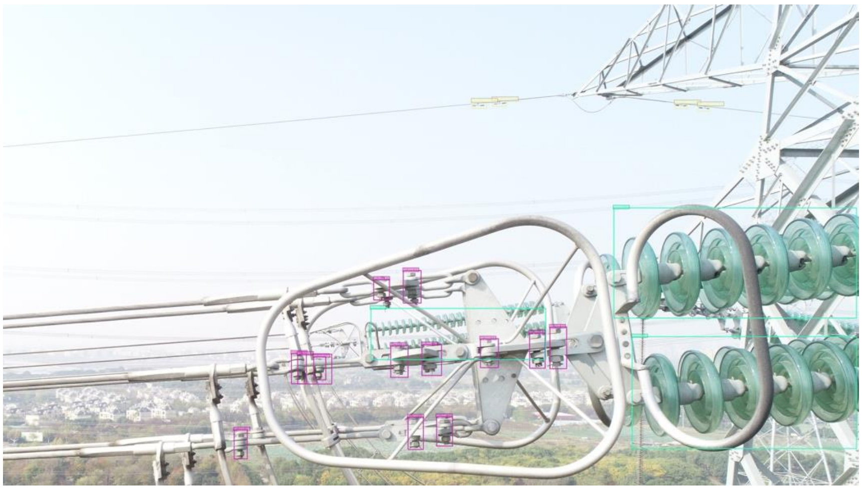

Figure 5. Detection result of an image with a cluster of pin bolts

Figure 4. Detection result on an aerial image with insulators of different scales and Figure 5. Detection result of an image with a cluster of pin bolts show the detection results of two aerial images. In 0, the output of plastic insulator, glass insulators, dogbone dampers and pin bolts are bounded respectively by blue, brown, magenta and cyan boxes. Undetected objects are bounded by red boxes. The trained model found all pin bolts, four glass insulators out of five and three dogbone dampers out of five, while classified an occluded damper as a dogbone damper at the bottom-left corner. In 0, glass insulators, dogbone dampers and pin bolts are bounded respectively by cyan, brown and magenta boxes. The trained model detected nearly every pin bolt in this dense scene except for two pin bolts end. Glass insulators in close distance are well bounded, and a string of

glass insulator occluded by pin bolts is detected, demonstrating that the trained model has properly learnt the features of glass insulators. Dogbone dampers in distance are all found.

From above analysis, it can be deduced that the recall rates of trained model on pin bolts and glass insulators are satisfactory to locate adequate pin bolts and insulators in an aerial image, revealing the potential of developing a cascaded fault detection system which takes detection result of SNIPER trained model as input to detect anomalies in pin bolts and insulators. Such cascaded fault detection system has been proposed in [11]. A head detection neural network is trained to locate insulators in an aerial image, then detected insulators are cropped and resized to be sent to the cascaded neural network which locates fault location.

4. Conclusion

In this work, a multi-scale training strategy named SNIPER is utilized to detect interesting power line equipment such as insulators, pin bolts and dampers. Trained model gives high recall rate and average precision on detection of pin bolts, insulators and dampers in the context of UAV based power transmission images. However, small objects as pin bolt faults and other views of pin bolts are not successfully trained probably due to their small size and similarity to other classes and imbalanced class distribution.

Future scope:

Add defects data of insulators, dampers and grading ring

Develop a cascaded fault detection system for pin bolts and insulators based on the high recall rate of SNIPER algorithm on the detection of pin bolts and insulators

Ameliorate the performance of detector in the case of imbalanced distribution of labelled data

Reference

[1] Nguyen V N, Jenssen R and Roverso D 2018 Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning International Journal of Electrical Power & Energy Systems 99 107–20

[2] Zhao Z and Cui Y 2018 Research progress of visual detection methods for transmission line key components based on deep learning Electric Power Science and Engineering 34 1–6

[3] Jian Zhang and Ruqing Yang 2006 Insulators Recognition for 220kv/330kv High-voltage Live-line Cleaning Robot 18th International Conference on Pattern Recognition (ICPR'06) 18th International Conference on Pattern Recognition (ICPR'06) vol 4 pp 630–3

[4] Zhao Z, Liu N and Wang L 2015 Localization of multiple insulators by orientation angle detection and binary shape prior knowledge IEEE Transactions on Dielectrics and Electrical Insulation 22 3421–8

[5] Huang J, Rathod V, Sun C, Zhu M, Korattikara A, Fathi A, Fischer I, Wojna Z, Song Y, Guadarrama S and Murphy K 2017 Speed/accuracy trade-offs for modern convolutional object detectors arXiv:1611.10012 [cs]

[6] Wang K, Wang J, Liu G, Zhou W and He Z RetinaNet Algorithm Based on Auxiliary Data for Intelligent Identification on Pin Defects Guangdong Electric Power

[7] Yang G, Sun C, Wang D, Jin T, Xu C, Lu Z and Zhang X 2020 Comparative Study of Transmission Line Component Detection Models Based on UAV Front End and SSD Algorithm Journal of Taiyuan University of Technology 51 212–9

[8] Singh B, Najibi M and Davis L S 2018 SNIPER: Efficient Multi-Scale Training arXiv:1805.09300 [cs]

[9] Girshick R 2015 Fast R-CNN arXiv:1504.08083 [cs]

[10] Girshick R, Donahue J, Darrell T and Malik J 2014 Rich feature hierarchies for accurate object detection and semantic segmentation arXiv:1311.2524 [cs]

[11] Tao X, Zhang D, Wang Z, Liu X, Zhang H and Xu D 2020 Detection of Power Line Insulator Defects Using Aerial Images Analyzed With Convolutional Neural Networks IEEE Transactions on Systems, Man, and Cybernetics: Systems 50 1486–98