A NOVEL SAMPLE LABELLING CRITERION FOR PIN DEFECT DETECTION IN UAV IMAGE

Kaidi Ying¹, Jianmei Wang^{1, *}

¹ College of Surveying and Geo-Informatics, Tongji University, Shanghai, 200092, China (2133693, 97031)@tongji.edu.cn

KEY WORDS: Pin Defect Detection, Convolutional Neural Network, Training Sample Set, Unmanned Aerial Vehicle Image, Object Detection, Power Line Inspection.

ABSTRACT:

With the rapid development of UAV technology, defect detection based on UAV images has expanded from power components such as insulators and dampers to bolts and pins. Different from the defect detection of insulators or dampers, there are two main difficulties in pin defect detection: (1) It is very small for bolts and pins compared to the entire image, usually less than 1%, and there are not enough features for detection; (2) Only bolts on link fittings need to be fixed with pins, while bolts in other parts do not need, so it is difficult to judge whether there are pin defects on bolts only based on the absence of pins on the bolts. Aiming at the above problems, cascade object detection method is adopted for pin defect detection in this paper, and improves the detection accuracy by gradually narrowing ROI (region of interest). The main contribution of this paper is to formulate a novel sample labelling criterion for cascade pin defect detection accuracy. In this paper, YOLOv4 is used to validate the proposed method. The result shows that compared with the existing sample set building methods, the proposed sample labelling criterion improves the accuracy from 85.2% to 92% and recall from 85.7% to 94.2%.

1. INTRODUCTION

UAV inspection of power equipment has become the main way of power industry inspection in China. With the rapid development of UAV technology, the objects of UAV inspection have expanded from defect detection of power components such as insulators and dampers to pin defect detection. As fasteners, pins and bolts are widely used in the connection among power components to stabilize the entire structure. If the pins fall off, the components will be loose and cause potential security issues. At present, the main research idea of pin defect detection is to transfer the general object detection model to the task of pin defect detection. For example, (Ning, 2019) used Faster R-CNN (Ren et al., 2017) to detect pin defects, and discussed the impact of different classifiers on the detection results; (Li et al., 2021b) replaced the feature extraction layer in Faster R-CNN with SCNet (Liu et al., 2020), which effectively improved the detection accuracy of pin defects; (Wang et al., 2020) implemented pin defect detection based on RetinaNet (Lin et al., 2017), and use GAN to improve the quality of training images. (Li et al., 2021a) improved the SSD (Liu et al., 2016) network structure to improve the detection accuracy of small objects. The above one-stage pin defect detection methods perform well in the images with a short shooting distance, but it is not effective for UAV images with a long distance and complex backgrounds.

Different from the defect detection of power components such as insulators and dampers, there are two main difficulties in pin defect detection on UAV images: (1) It is very small for bolts and pins compared to the entire image, usually less than 1%, and there are not enough features for direct object detection to determine whether pins are missing; (2) Only the bolts on the link fittings need to be fixed with pins, while the bolts in other parts do not need, so it is difficult to judge whether there are pin defects on bolts only based on the absence of pins on the bolts. Aiming at the above problems, a cascade pin defect detection method has been proposed as shown in Figure 1. This method first uses the connection part detection model to extract the connection parts in the original images and generate a series of connection part images, and then inputs these images into the pin defect detection model to identify bolts with pins missing.

Obviously, cascade object detection is a serial process, and the detection accuracy of pin defect is highly dependent on the definition of the connection parts at the first stage. (Wang et al., 2021) defined the connection parts as insulator connection lines, insulator connection towers and line connection towers, and images of connection parts extracted from such annotations can only partially filter out bolts that do not need pins. (Xu et al., 2020) defined the connection parts as electric power fittings such as shackles, triangular plates, adjustment plates, and suspension clamps. The narrowed ROI (region of interest) can filter out most of the bolts that do not need pins and irrelevant backgrounds. It also solves the problem that pins and bolts are too small in the images.

However, there are still two problems with the two methods mentioned above. One is that the diversity and complexity of connection parts lead to high cost of building sample sets and low detection accuracy. And the other is about overlap of the bounding boxes labelling adjacent connection parts samples, as shown in Figure 3. Based on such training samples, not only the CNN model converges slowly, but also the trained model has a large number of missed detections during actual detection, so that the ROI of pin defects cannot be input into the secondstage object detection.

^{*} Corresponding author

Aiming at the above problems, a new labelling criterion for training samples of pin defects is proposed in this paper. Link fittings are defined as connection parts for the first-stage object detection. And by improving the definition of the ROI of bolts and pins, the accuracy of the cascade pin defect detection method is effectively improved. The remainder of the paper is organized as follows. Section 2 details the construction of training sample sets for cascade pin defect detection. Section 3 briefly introduces the object detection algorithm of YOLO (Redmon et al., 2016) for experiments. Section 4 is experimental results and analysis. The last section gives some conclusions and future work.



Figure 1. Pipeline of the cascade pin defect detection.

2. TRAINING SAMPLE SETS FOR CASCADE PIN DEFECT DETECTION

2.1 The Proposed Link Fitting Dataset

Link fittings refer to fittings that combine insulators, clamps, and protective fittings into suspension or tension strings through bolted connections, so there must be bolt-pin assemblies on them. Since the first-stage detection significantly affects the accuracy of the cascade detection, the quality of link fitting dataset is quite important. However, not all link fittings can be chosen as detection object. On the one hand, the reduction of object types reduces the cost of building sample sets. On the other, link fittings are hooked to each other, resulting in lots of overlap or even nesting of the bounding boxes when labelling adjacent link fittings, which directly influence the quality of link fitting dataset.

Therefore, eight specific types of link fittings are selected as detection object. Their basic shapes are shown in Figure 2, and from (a) to (h) are respectively ball eye, socket-clevis eye, U-bolt, shackle, eye chain links, clevis, yoke plate (type P), and yoke plate (type PS).



Figure 2. Link fittings: (a) ball eye, (b) socket-clevis eye, (c) Ubolt, (d) shackle, (e) eye chain links, (f) clevis, (g) yoke plate (type P), and (h) yoke plate (type PS).

Figure 3 shows the comparison of our method with the existing methods for labelling detection objects in images. For each pair of images, the magenta boxes label all kinds of fittings on the left, and the orange boxes label eight specific link fittings mentioned before on the right. Through comparison, it can be seen that the method proposed in this paper not only reduces the

object ROIs significantly, which further increase the area ratio of bolts and pins during subsequent pin defect detection, but also decreases the overlap between the labelling bounding boxes. Meanwhile, with the reduction of the types of link fittings, both the labelling cost and the ambiguity of different manual labelling samples are relatively reduced.



Figure 3. Comparison of four pairs of the images labelled with different detection objects. Column (a) is the existing methods using the magenta boxes to label all kinds of link fittings, and column (b) is our method using the orange boxes to label eight specific link fittings.

2.2 Bolt-Pin Assembly Dataset

At second stage for pin defect detection, it is necessary to construct a bolt-pin assembly dataset. First, obtain the image data. Because the second stage is to detect the bolt-pin assemblies on the link fittings, all the images containing link fittings at the first stage need to be cropped out as the image file of this dataset.

Second, label the samples in the dataset. The purpose of pin defect detection is to judge whether the pins are missing, so the dataset is divided into two classes, bolts with pins and bolts with pins missing. However, since the shape of bolt head is similar to the above two classes, it is also divided into another class to prevent misdetection. Figure 4 shows some examples of three classes, the orange boxes in column (a) show normal bolts with pins, the magenta boxes in column (b) contain bolts with pins missing, and the blue boxes in column (c) are bolt heads on the front side of the bolt.



Figure 4. Three classes of bolt-pin assemblies, the orange boxes in column (a) are normal bolts with pins, the magenta boxes in column (b) show bolts with pins missing, and the blue boxes in column (c) contain bolt heads.

3. CASCADE PIN DEFECT DETECTION ALGORITHM: YOLO

There are a lot of detection algorithms that can be used for pin defect detection, and YOLO is used as an example method for experiments in this paper, which does not have an impact on the following conclusions.

YOLO is called you only look once. As the name says, its detection speed is very fast. Just after taking a look, the result can be detected. The YOLO algorithm redefined object detection as a single regression problem. Its core idea is to use the entire image as the input layer of the network, and directly return the position of the bounding box and its category in the output layer of the network.

The development of YOLO series so far mainly includes v1, v2, v3, v4 and v5 and the improved series for each version. YOLOv1, proposed by Redmon is a one-stage detection framework for real-time high-performance object detection. The position and category of the object in the image can be predicted by only inputting the image into the network once. It has the advantages of fast speed and easy deployment, but the object positioning accuracy is poor, especially for small objects.

YOLOv2 (Redmon and Farhadi, 2017) improved YOLOv1 from three aspects: predict faster, more accurate, and recognize more objects. It redesigned the backbone network Darknet-19 for feature extraction, and added batch normalization. Also, it introduced anchor boxed from Faster R-CNN and used Kmeans to cluster the prior anchor. It improved the positioning accuracy of objects but there were still problems in small object detection.

YOLOv3 (Farhadi and Redmon, 2018) optimized the model on the basis of YOLOv2. By drawing on the idea of ResNet, it used Darknet-53 to solve the problem of gradient disappearance or gradient explosion. And, it introduced FPN (feature pyramid network) for multi-scale prediction, which effectively improved the result of small object detection.

YOLOv4 (Bochkovskiy et al., 2020) combined lots of previous research techniques, added many practical tricks and make appropriate innovations. Based on YOLOv3, it combined and tested over twenty tricks in the field of object detection, and achieved a balance between detection speed and accuracy.

YOLOv5 (Jocher et al., 2021) further optimized the YOLO series models, and improved the detection performance by adding Focus structure, adaptive image scaling and other methods.

YOLOv4 model is chosen as an example for the following experiments in this paper, as shown in Figure 5, because it is a mature and effective version of YOLO and is widely used in many real-time detection tasks. It should be noted that with the rapid development of object detection, there will be more advanced algorithms available for the sample labelling criterion of pin defect detection proposed in this paper.



Figure 5. YOLOv4 model structure.

4. EXPERIMENTS RESULTS AND DISCUSSION

4.1 Implementation Details

The experimental platform in this paper is configured as follows: the CPU is Intel(R) Core(TM) i9-9900K, the GPU is NVIDA RTX 2080Ti with 11GB memory, and the RAM is 16GB; the operating system is Ubuntu 20.04, and the implementation framework is Pytorch.

The experimental data was from China Southern Power Grid Co., Ltd. It is RGB images captured by multirotors from multiple angles near the transmission line. The image size is 4000×3000 pixels. The image quality is high and the objects are clearly visible in most images.

The fittings dataset A and the link fittings dataset B are built respectively by the existing method and the method proposed in this paper. The corresponding bolt-pin assembly datasets C and D are also built. Dataset A contains 2,600 pictures with 3,944 fitting objects while dataset B contains 2,172 pictures with 4,786 link fitting objects. Dataset C contains 2026 pictures, including 3139 bolts with pins and 349 bolts with pins missing; dataset D contains 3724 pictures, including 4342 bolts with pins, 509 bolts with pins missing and 196 bolt heads.

Four YOLOv4 models are trained separately with the above datasets. The training adopts fine tune with freeze training method, using the pre-training model yolov4_weight.pth as the initial weight. Set the initial learning rate = 0.0001, batch size = 4. Freeze the shallow network to train 50 epochs and then unfreeze the network to train another 50 epochs.

4.2 Results and Discussion

For the first stage detection, the most important thing is whether the objects are missed. It means whether the recall can be as close to 100% as possible. Therefore, at this stage the recall is the main criterion for experimental evaluation and AP (average precision) is the second.

Table 1 shows that the recall and AP of dataset B are much higher than those of dataset A. The results prove that the method proposed in this paper is superior to the existing method. By choosing smaller and more specific link fittings as the detection object, the quality of dataset is higher and the trained model can predict more accurately.

Detector	Dataset	Epoch	Batch_size	Recall	AP
YOLOv4	А	100	4	81.9%	0.830
	В			94.9%	0.937

Table 1. Comparison of recall and AP of dataset A and Brespectively detected by YOLOv4.

As shown in Figures 6, use YOLOv4 models to test datasets A and B, and get the detection results. Columns (a) and (b) show the comparison between the results predicted and the ground truth by the existing method, and columns (c) and (d) show the comparison between the results predicted and the ground truth by the method in this paper. Obviously, there are a great number of missed detections with the existing method. For example, no fittings are detected in row 1. Secondly, the object position accuracy is low. In rows 2, 3 some detection boxes are quite different from the ground truth. In addition, it has limitations in labelling and is easy to miss bolts that are not on normal fittings, such as last row. But the detection results using the method in this paper are close to the ground truth, and there are almost no missed objects, and the bolts to be detected at the second stage are nearly included in the bounding boxes.



Figures 6. Use YOLOv4 models to test datasets A and B, and get the detection results. Columns (a) and (b) show the comparison between the results predicted and the ground truth by the existing method, and columns (c) and (d) show the comparison between the results predicted and the ground truth by the method in this paper.

On the one hand, the objects selected by the existing method have various types and different shapes from different shooting angles. When the dataset is small, the model does not learn enough features for small samples, resulting in misdetection. On the other hand, when the existing method labelled the objects, the bounding boxes of some adjacent fittings were not accurate enough. It affected the positioning accuracy during target detection, and lead to missed or false detection of the fittings.

However, the specific link fittings are selected instead of all the fittings as detection objects. It is simpler and ensures the bolts to be detected must be in the region of interest, so that no objects will be missed. Moreover, the shapes of link fittings at different angles are similar and the objects overlaps less. It reduces the difficulty of the detection tasks and makes the missed detection less and position more accurate.

Table 2 shows the detection results of bolts with pins and bolts with pins missing in datasets C and D. Both methods to detect bolts with pins perform well. But when detecting bolts with pins missing, our method is higher than the existing method in terms of precision, recall and AP with, especially the recall is increased to 94.2%. It greatly reduces the probability of missed detection.

Figure 7 reveals the results of pin defects detection in the UAV images by using the method in this paper. Picture (a) is the output detection result of UAV images, and picture (b) zooms in the orange box in picture (a). The red boxes draw the bolt with pin missing, which needs attention.

Detector	Dataset	Class	Precision	Recall	AP
YOLOv4	С	bolt with pin	96.6%	97.6%	0.985
		bolt with pin missing	78.1%	85.7%	0.853
	D	bolt with pin	96.1%	97.8%	0.971
		bolt with pin missing	85.1%	94.2%	0.92





(a)

(b)

Figure 7. Results of pin defects detection in the UAV images by using the method in this paper. (a) is the output detection result of UAV images, and (b) zooms in the orange box in (a).

5. CONCLUSION

Object detection method based on deep learning is the main method for UAV pin defect detection. The accuracy of deep learning-based object detection is strongly dependent on the quality of the training sample set. Based on the characteristics of bolts and pins in UAV images, a novel sample labelling criterion is proposed in this paper. Not only is it less work to build a sample set according to this criterion, but the detection accuracy of the model trained based on this sample set is higher. Therefore, the cascaded pin defect detection method and the sample labelling criterion proposed in this paper are of great practicality.

Moreover, there is usually an imbalance of positive and negative samples in defect detection. It is the same for pin defect detection. Therefore, according to the features of bolts and pins on power transmission lines, designing better data augmentation methods will be our research content in the future.

ACKNOWLEDGEMENTS

This work is supported by National Key R&D Program of China (2018YFB0505400).

REFERENCES

Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y.M., 2020. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:*, 2004. 10934.

Farhadi, A., Redmon, J., 2018. Yolov3: An incremental improvement. *Computer Vision and Pattern Recognition*, 1804.

Jocher, G., Stoken, A., Borovec, J., et al., 2021. ultralytics/yolov5: v5.0 - YOLOv5-P6 1280 models. Zenodo. doi.org/10.5281/zenodo.4679653.

Li, R., Zhang, Y., Zhai, D., et al., 2021a. Pin Defect Detection of Transmission Line Based on Improved SSD. *High Voltage Engineering*, 47, 3795-3802.

Li, X., Liu, H., Liu, G., et al., 2021b. Transmission Line Pin Defect Detection Based on Deep Learning. *Power System Technology*, 45, 2988-2995.

Lin, T.Y., Goyal, P., Girshick, R., et al., 2017. Focal Loss for Dense Object Detection. 2017 IEEE International Conference on Computer Vision (ICCV), 2999-3007.

Liu, J.J., Hou, Q., Cheng, M.M., et al., 2020. Improving Convolutional Networks With Self-Calibrated Convolutions. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 10093-10102.

Liu, W., Anguelov, D., Erhan, D., et al., 2016. SSD: Single Shot MultiBox Detector. *Computer Vision – ECCV 2016*, 21-37.

Ning, B., 2019. Pin fault detection in UAV inspection image based on deep learning. *Computer Measurement & Control*, 27, 25-29.

Redmon, J., Divvala, S., Girshick, R., et al., 2016. You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788.

Redmon, J., Farhadi, A., 2017. YOLO9000: Better, Faster, Stronger. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 6517-6525.

Ren, S., He, K., Girshick, R., et al., 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, 1137-1149.

Wang, H., Zhai, X., Chen, Y., 2021. Two-stage pin defect detection model based on improved Cascade R-CNN. *Science Technology and Engineering*, 21, 6373-6379.

Wang, J., Wang, K., Liu, G., et al., 2020. Recognition of defects in pins based on generative adversarial network and RetinaNet. *Journal of South China University of Technology*(*Natural Science Edition*), 48, 1-8.

Xu, W., Wang, J., Zhang, S., et al., 2020. Missing Pins Detection in Power Transmission Line using Unmanned Aerial Vehicle Images. *The 7th China High Resolution Earth Observation Conference*, 1-11.