ISBN: 978-981-18-7950-0
2023 the 13th International Workshop on Computer Science and Engineering (WCSE 2023)
doi: 10.18178/wcse.2023.06.034

RGB-W based Hazard Detection in Electric Power Communication Networks

Chengrun Qiu [1 +], Wei Guan[2], Yuanhang Huang [1] and Zefeng Jia [1]

[1] State Grid Wuxi Power Supply Company, Wuxi, China

[2] State Grid Yancheng Power Supply Company, Yancheng, China

Abstract. Hazard detection is an important computer vision area in electric power communication network. In this paper, we propose a hazard detection method in electric power communication network. This method turns the short-range wireless localization information into a confidence mask corresponding to the image size. The mask is able to be input into object detectors to revise confidence scores. The proposed method aims to increase the accuracy of the detections and alleviate the problems of false positives and false negatives. We design two kinds of confidence mask in order to find a better method to utilize the localization information. In addition, dataset containing specific scenes is collected to train and test the proposed method of hazard detection task in electric power communication network. Finally, this paper adopts various metrics to evaluate the effectiveness of the proposed method. The experiments demonstrate the feasibility and superiority of the proposed method in both theory and practice.

Keywords: hazard detection, electric power communication network, wireless localization, amplitude heatmap, confidence mask.

1. Introduction

With the development of information technology and computing device, the demand for artificial intelligence in electric power area has become more and more popular. How to apply the computer vision or big data methods to the problems in electric power area turns to be a popular topic. At present, the electric power communication network uses optical cable technology, and most of the communication cables are ordinary optical cables, OPGW optical cables, and ADSS optical cables. Among them, ordinary optical cables and ADSS optical cables are non-metallic optical cables, which face a risk of being damaged by external forces from external hazard. The hazard detection is an important issue to ensure the smooth operation of electric power communication network, which has a more strict application standard than industry or civil areas.

Object detection is an important area in computer vision, which can be applied in the tasks of hazard detection in communication optical cables. The state-of-the-art detection methods can achieve sufficient accuracy in industry and civil field [1], [2]. The state-of-the-art detectors based on neural network backbones can perform well on current representative official datasets, e.g. Microsoft COCO [3]. These methods can get high scores in mean average precision (mAP) proposed in Microsoft COCO, which is the most widely used metric in object detection. Some improved detection frameworks have been conducted into industry areas like indoor scenes [4] and vehicle detection [5]. Generally, most of the detectors under specific scenes are born from the representative detectors, including the R-CNN series [6], [7] and the one-stage series such as YOLO [2]. For R-CNN series, the development of these detectors focuses on various areas, including providing sparse proposals [8], extracting features with more scales [9], proposing dynamic convolution backbones [10] and adpoting balance training [11]. Many literature now focuses on one-stage detector. Among the one-stage detectors, YOLO [2] series are the most famous and have developed into the official version YOLOv5 so far. The detector's backbone, decoupled head, training label assignment and training loss function have evolved dramatically.

+ Corresponding author. Tel.: +86-134-3889-0825.
E-mail address: qiucr@js.sgcc.com.cn.

However, the state-of-the-art detectors will still face problems that the detector's performance can't satisfy the specific standard in electric power area. With the physical assistance of the wireless signal's localization, the disadvantage of single image detection can be reduced. Some literature attempts to combine wireless information with image data for detection or localization tasks. In [12], a deep learning parking detection algorithm based on W -shape magnetic wireless sensor network is proposed for large parking lots with vertical parking spaces. For identity identification, [13] uses the YOLOv2 model to process the visible-light eye images marked with the jointly partial iris and sclera region. For post COVID-19 preventive measures, [14] proposes an IoT-based architecture including the deep learning technologies and the body temperature monitor.

In this paper, a detection framework is proposed which inputs the short-range wireless localization confidence mask as a bypass. The method converts the localization information into a confidence mask, including the arrive of angle (AoA) and the time of flight (ToF). This paper designs two kinds of confidence mask, one is the dense mask and another is the grid mask. Then the method aligns the confidence mask with the image's size and the camera's effective focal length to form an initial estimate of each moving object. Finally, in order to increase the detector's accuracy, the bounding box's confidence score will be corrected by the confidence mask before non-max suppression. In order to verify the feasibility of the method proposed in this paper, the corresponding training and test dataset for hazard detection in electric power communication network is collected. The performance metrics adopted in the dataset include mean average precision (mAP), average false positives and false negatives per image (FP\&FN per image), and true detection ratio. With sufficient metrics, the experiment results of our dataset show that the proposed method in this paper can improve the accuracy of the hazard detection task and improve visual experience.

2. Localization-detection Fused Method

2.1. Wireless localization

The proposed method in our paper doesn't limit the implementation of the localization, but only requires the localization error less than 0.5 meter. Considering the realizability of the proposed method in electric power communication network's monitor, we emphasize the localization to be short-range and the localization range is lower than 10 meter from the antenna. In this case, we choose the localization method based on AoA-ToF joint estimation in single access point near the image monitor. The access point is equipped with vertical and horizontal antenna arrays, which can measure and calculate the horizontal angle, vertical angle and distance of the moving objects.

2.2. Detector framework

The proposed method in our paper doesn't limit the implementation of the detection, but only requires that the detector works based on the bounding boxes output together with the confidence scores. Although the performance of existing detection methods can satisfy ordinary detection requirements, there still exists problems of false positives and false negatives, which is a hidden damage to the electric power communication network. This article converts short-range wireless localization information into a confidence mask, which serves as a bypass input into the detection framework and provides a physical assistance for improving detection accuracy. The detection framework proposed in this article that uses localization information as a bypass input is shown in Figure. 1.

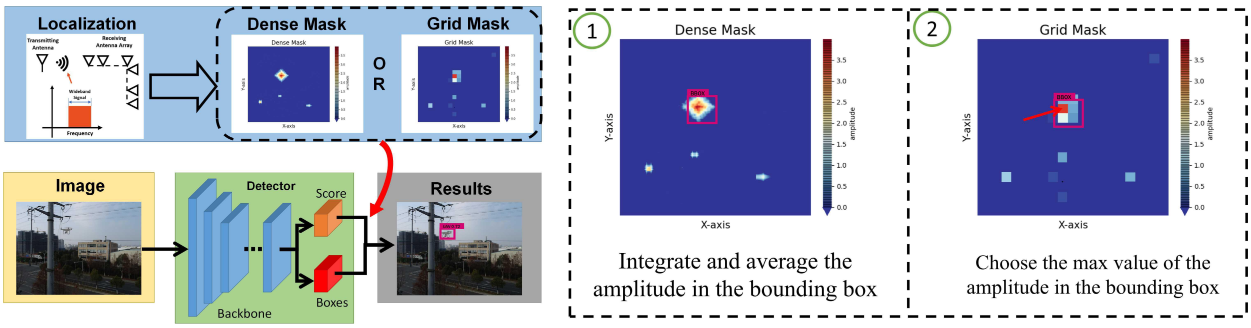

Fig. 1: Hazard detection framework with the localization-based confidence mask bypass input.

We use confidence masks transferred from short-range localization to adjust the confidence score of each bounding box before non-max suppression. In this paper, we propose two ways to transfer the localization information into confidence mask. One way is to simply utilize and normalize the localization amplitude and make it into a confidence mask, which is called dense mask. The other is to divide the dense mask to grids and normalize the amplitude of each grid, which is called grid mask. When utilizing dense mask, the framework will integrate the localization amplitude of dense masks in each bounding boxes as original correction factor γ^0. The method of obtaining γ^0 is shown on the right side of Figure. 1.Then assuming there exist N bounding boxes, we work out the decay factors γ of the corresponding bounding boxes as:

$$\{\gamma_1, \gamma_2, ...\gamma_N\} = \left\{\frac{\gamma_1^0}{\tau}, \frac{\gamma_2^0}{\tau}, ..., \frac{\gamma_N^0}{\tau}\right\}, \left(\tau = \max\{\gamma_1^0, \gamma_2^0, ..., \gamma_N^0\}\right). \tag{1}$$

Different from the dense mask, the grid mask determines the bounding box's original correction factor by the max localization amplitudes of the grids included fully or partly in the bounding box. Then the decay factor γ can be obtained by equation (1). Finally we get the confidence score of each bounding box their as below:

$$S_{\text{new}} = (1 - \lambda + \lambda \times \gamma) \times S \tag{2}$$

The factor λ is a pre-defined constant factor based on the expected accuracy of the applied localization method. After processed by the confidence mask, the bounding boxes with the localization amplitudes' authentication will have more confidence scores than those of the boxes deviating from any localization amplitude pixels.

2.3. Electric power communication network dataset

The category of the detection object is determined by the training dataset, and the detection framework proposed above is lack of specific training. Due to the confidentiality of electric power communication network, there is no public dataset for hazard detection task at present. It is necessary to collect images or videos in specific scenes of the electric power communication network as dataset. The proposed method filters the categories of the existing object detection dataset to pre-train the hazard detector to classify the specific categories. The pre-training dataset comes from Microsoft COCO dataset and Det-Fly dataset [14]. The pre-training focuses on 12 categories of hazard, including UAVs, various animals, digging trucks and other objects.

Then, this paper constructs a hazard detection dataset in electric power communication network. In order to collect the dataset, this article synchronizes image and localization data to collect a dataset composed of image and localization information, i.e. RGB-W information. The training sample for this dataset contains about 5000 frames of image together with localization information, and the validation sample contains about 1000 frames of image together with localization information. In order to increase the diversity in specific situations, we collected a portion of image data with fewer brightness at the scene of evening and overcast sky. Some frames in the dataset may not include any hazard to help train and reduce the false positive of the detection framework.

3. Experiment Results

The two-stage detection methods used in our experiment are Mask R-CNN and Sparse R-CNN with the ResNext-101-32x4d backbone. The one-stage detection method used in our experiment is YOLOv5, with the the official archive's backbone model weights. The adapted backbone for YOLOv5 is CSPDarkNet and the image will be resize into 608 × 608 pixels before input. Based on official archives, the confidence score thresholds for Mask R-CNN, Sparse R-CNN and YOLOv5 are all 0.05 and the intersection over union (IoU) thresholds are all 0.5.

The first metric we adopt in our experiment is the mean average precision(mAP) defined in COCO, which is the most widely used metric in object detection area. In this paper, two additional evaluation metrics have also been used in our collected dataset. One of the adopted evaluation metric is the average number of false positives and false negatives in each image (FP\&FN per image). The other adopted metric is called true detection ratio. Different from mAP, these two metrics can demonstrate the advantages of the proposed method in visual results. The evaluation method in our dataset focuses more on the visual experience of hazard detection in electric power communication network. The evaluation method counts all detection bounding boxes obtained by confidence threshold filter. Specifically, the method doesn't need to rank the bounding boxes with their confidence scores in descending order.

Table 1: Comparison between methods with/without localization mask.

Methods	**Localization Mask**	**AP**	**AP$_{50}$**	**AP$_{75}$**	**FP & FN per image**	**True Detections Ratio(%)**
Mask R-CNN	None	60.9	77.5	62.4	0.792	76.47
	Dense	62.4	80.3	64.8	0.608	81.25
	Grid	**63.9**	**82.4**	**66.5**	**0.595**	**82.67**
Sparese R-CNN	None	63.6	76.8	64.7	0.704	78.48
	Dense	65.8	**79.5**	66.2	0.585	82.52
	Grid	**66.8**	78.8	**68.4**	**0.569**	**83.94**
YOLOv5	None	55.5	74.6	58.8	0.399	84.69
	Dense	60.9	76.8	60.2	0.223	86.83
	Grid	**64.9**	**78.4**	**66.0**	**0.187**	**89.75**

Table 1 shows the performance improvement of the proposed detection framework with the localization bypass input on the dataset, with the 0.05 confidence thresholds for all detection methods. The results in show that by using localization-based confidence mask input, the AP values of the three detection methods can be significantly improved. The results also show that with the use of the localization-based confidence mask input, our proposed method can significantly reduce the number of false positives and false negatives per image, and improve the ratio of correct bounding boxes. Moreover, the performance of grid mask is a little better than that of dense mask even though dense mask contains more amplitude information. This is because if some slight shaking occurs in localization, the dense mask will reflect its amplitude in detail and change the integration in each bounding boxes. The grid mask can tolerate slight localization shaking because the mask's grid will still cover the max amplitude in some cases.

Figure. 2 shows some visible results of the scenes in our collected detection dataset. The first column shows the comparison results of utilizing dense mask or not in YOLOv5, while the second column shows the comparison results of utilizing grid mask or not in YOLOv5. Different from the official archive, we increase the confidence threshold score to 0.5 in YOLOv5 to decrease the number of false positives. The image results indicate that compared with single detector YOLOv5, our proposed detection framework with localization bypass input can effectively increase the accuracy and improve the observer's visual experience. On the other hand, with either utilizing dense mask or grid mask, there doesn't exist an obvious difference between the visible results.

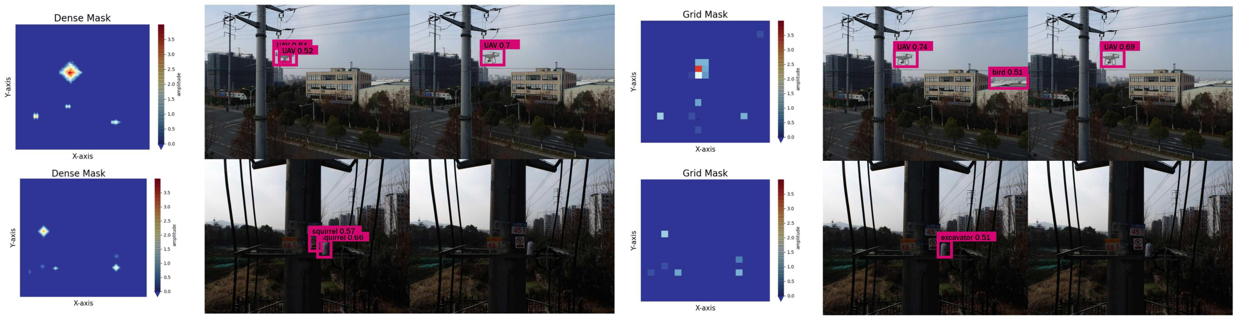

Fig. 2: Detection results in real-world datasets.

4. Conclusion

This paper proposes a method to input the short-range wireless localization information into the object detection framework to improve the accuracy of the hazard detection in electric power communication network. This paper first uses short-range wireless localization methods to obtain the angle and distance data from the moving hazard to the camera through the characteristics of the reflected signal. Then, this paper proposes a method that the short-range wireless localization mask is used as a bypass input to the hazard detection framework for both one-stage detection frameworks and two-stage detection frameworks. Specifically, this paper designs both dense mask and grid mask as a bypass and input the localization information into the detection framework. After trained with the categories of hazard to be detected and trained with the dataset of the collected scenes in electric power communication network, the experiment results on the validation dataset show that the proposed method can improve the accuracy of the hazard detection task and can improve visual experience by reduce false positives and false negatives.

5. References

[1] K. He, G. Gkioxari, P. Doll´ar, and R. Girshick, "Mask R-CNN," *IEEE Trans. Pattern Anal. Machine Intell.*, 2020, 42 (02) :386–397.

[2] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA: 2017.

[3] T. Y. Lin, M. Maire, S. Belongie, et al., "Microsoft coco: Common objects in context," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus *(CVPR)*, OH, USA, 2014.

[4] J. Ni, K. Shen, Y. Chen and S. X. Yang, "An Improved SSD-Like Deep Network-Based Object Detection Method for Indoor Scenes," *IEEE Trans. Instrum. Meas.*, 2023, 72: 1-15.

[5] Y. Ding, Y. Qu, D. Du, et al., "Long-Distance Vehicle Dynamic Detection and Positioning Based on Gm-APD Lidar and LIDAR-YOLO," *IEEE Sensors J.*, 2022, 22(17): 17113-17125.

[6] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Region-based convo lutional networks for accurate object detection and segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, 2016, 38(01): 142-158.

[7] S. Ren, K. He, R. Girshick and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Machine Intell.*, 2017, 39(06): 1137-1149.

[8] P. Sun, R. Zhang, Y. Jiang, et al., "Sparse R-CNN: End-to-End Object Detection with Learnable Proposals," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, pp. 14449–14458, 2021.

[9] J. Li, X. Liang, S. Shen, et al, "Scale-aware fast r-cnn for pedestrian detection," *IEEE Trans. Multimedia*, 2018, 20(04): 985-996.

[10] Q. Hong, F. Liu, D. Li, et al., "Dynamic Sparse R-CNN," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, pp. 4713–4722, 2022.

[11] H. Qiu, H. Li, Q. Wu, et al, "Hierarchical cntext features embedding for object detection," *IEEE Trans. Multimedia*, 2020, 22(12): 3039-3050.

[12] Y. Zhang, C. Wu, P. Wu, Y. You and Y. Yuan, "A Deep Learning Parking Detection Algorithm Using W-Shape Magnetic Wireless Sensor Networks," *IEEE Trans. Instrum. Meas.*, 2023, 72: 1–13.

[13] C.W. Chuang and C.P. Fan, "Deep-Learning Based Joint Iris and Sclera Recognition with YOLO Network for Identity Identification," *Journal of Advances in Information Technology*, 2021, 12(01): 60-65.

[14] A. Akib, K. Nur, S. Saha, et al., "Computer Vision-Based IoT Architecture for Post COVID-19 Preventive Measures," *Journal of Advances in Information Technology*, 2023, 14(01): 7-19.

[15] Z. Ye, Z. Chen, D. Lv, et al., "Air-to-air visual detection of micro-uavs: An experimental evaluation of deep learning," *IEEE Robot Autom. Lett.*, 2021, 6(02): 1020–1027.