Multi-object Detection by Using CNN for Power Transmission Line Inspection

Dinh Cong Nguyen[1,2](✉), The Cuong Nguyen[1,2], Dinh Hung Phan[2], Nhan Tam Le[3], and Van Vien Tran[2]

[1] Hong Duc University, Thanh Hoa, Vietnam
{nguyendinhcong,nguyenthecuong}@hdu.edu.vn
[2] ThinkLABs JSC, Thanh Hoa, Vietnam
{hung,vientv}@thinklabs.com.vn
[3] IBM Vietnam, Hanoi, Vietnam
letam@vn.ibm.com

Abstract. Multi-object detection for power transmission line is one of the key tasks to control and monitor quality of the system. In the past, defective objects were found out by naked-eye inspection through aerial images and relied on the experienced workers. Due to harsh environmental conditions, manual observation might be a time-consuming and dangerous task. Recently, this task has been supported by machine learning where deep-learning algorithms are applied to increase the efficiency of detection/recognition phases. This paper discusses different approaches for multi-object detection based on Convolutional Neural Network (CNN) model to investigate the quality and condition of power lines in Vietnam. Our proposed system outperforms the state-of-the-art methods on our dataset.

Keywords: Power transmission inspection · Multi-object detection · Aerial image

1 Introduction

Power transmission system is composed of different types of components such as tower, insulator, conductor, so on. These components play a critical role in the safe transmission of the electric power system. Because these components endure extreme weather conditions (raining, storm, etc.), they have sometime suffered several damages such as: insulator faults, connectors corroded, conductor damaged Fig. 1. Moreover, material aging, overloading are also main factors to influence on them. The quality of the system is whereby decreased.

Therefore, to make the power supply system sustainable, reliable and available, these components need to periodic inspection throughout the year. This process helps to detect defective devices. Then, the experts will decide which parts should be repaired or replaced by the new ones. Previously, the conventional inspection was mostly based on a manual inspection. However, this task

N.-S. Vo et al. (Eds.): INISCOM 2021, LNICST 379, pp. 337–347, 2021.
https://doi.org/10.1007/978-3-030-77424-0_28

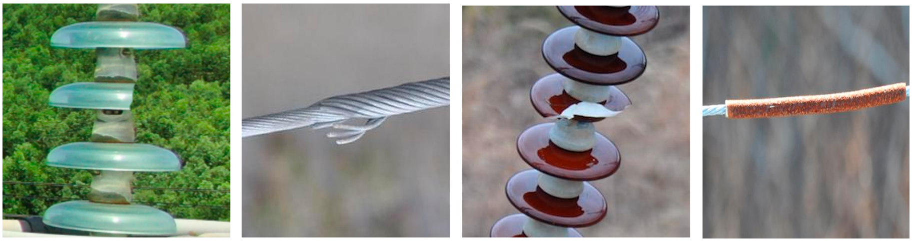

Fig. 1. Due to outdoor environment in complicate land forms and unpredictable weather, the power transmission system could be damaged frequently.

remains a challenge because of time-consuming, costly and life-threatening processes. Due to these reasons, it needs to find out an efficient method to cope with this problem. An automatic observation is an alternative solution [1], known as multi-object detection and recognition in the literature.

Recently, the development of Unmanned Aerial Vehicle (UAV) and new digital image processing technologies could support to capture images or videos with very high resolutions at different angles and distances. A new platform for power transmission inspection has been supported [2]. This process could be divided into two phases as data collection and data analysis. Although UAV inspection greatly decreases in works of inspectors, low cost, high security, efficiency, but it also creates enormous data. Moreover, a large number of transmission towers are built in the forests or mountains. Therefore, the collected aerial images are suffered from complex backgrounds, illumination changes, rotations, distances that directly impact on the quality of detection and recognition.

For our context, we collected images and videos from the power transmission line at The National Power Transmission Corporation (EVN NPT PTC2)[1] with different scenarios in order to detect those objects in this system. Therefore, we need to study several approaches to find out the best model to fit well with our own dataset. Recently, Convolutional Neural Network (CNN)-based detection approach has gained big achievements in terms of a very high detection accuracy and a strong robustness with gigantic datasets [3]. This highly motivates us to apply CNN model for multi-object detection of our problem.

In this paper, a new approach to automatically recognize different elements of power transmission devices is given. The main contributions are follows.

- A new standard dataset is produced by collected and manually labeled images from EVN NPT in Vietnam.
- A new system to automatically detect five types of components in the aerial images is proposed by adopting a deep convolutional neural network.

The remainder of this paper is given as follows. Section 2 revises several works related to this topic. Then, Sect. 3 describes our dataset. Next, Sect. 4 details multi-object detection tasks while Sect. 5 will present our system. Section 6 shows

[1] https://www.npt.com.vn.

and discusses the results. At last, Sect. 7 will conclude and discuss some perspectives.

2 Related Works

The existing solutions to deal with this task could be divided into two categories: hand-crafted feature-based approaches and deep learning-based approaches. In hand-crafted feature-based approaches, a variety of low-level features are applied such as color [4,5], shape [7], edge [8], texture [9] and so on. However, as summary from [10], the methods are less accurate in average. Basically, they are quite sensitive with noises, complex backgrounds. In addition, their performances have degraded in case of which electric devices are overlapped or by filming distances.

Over past few years, along with the development of hardware devices and deep neural network models, object detection and recognition have obtained a great achievement. Many of works were adopted convolutional neural network (CNN) model into the inspection of power lines. It could be listed in [3]. As [11] applied CNN into their work for detection process and the color/edge-based models for segmentation process. Although this contribution is robust to object detection, its effectiveness limits with a small number of scenarios due to the sensitiveness of low-level features. For insulator detection, [12,13] are applied either R-CNN or Faster R-CNN model. These works divided detection process into two steps in which the first step localized insulators and then the second step classified defective insulators. The main drawback is that two steps are strongly dependent. It results in poor classification if the first step does not perform well. Moreover, it could be hard to train an end-to-end system with a reasonable performance for the insulator fault detection due to complex backgrounds and various angles of shooting. To address this problem, [10,14] employed a CNN architecture and used either the segmentation of the insulator fault or the adaptive morphology method. However, these approaches are quite sensitive with the complex and uncontrollable scenarios, because they applied low-level feature-based model after the detection for classification.

It can be seen that most of the recent works on this topic employed two-stage strategy where CNN was applied in the first stage to detect objected, followed by segmentation with low-level features to classify whether faulty devices or not. As known that low-level features are quite sensitive, they therefore cannot be robust against the occlusion, various orientations, illumination changes, viewpoint changes, and complex background. Our approach differs from those, we consider faulty devices as individual objects, and hence, we just employ one-stage model.

3 Data Collection

At the best of our knowledge, there is no publicly available dataset that identifies the devices on the high-voltage transmission lines in Viet Nam. To make it, we have created a dataset called "ThinkLabs-data". ThinkLabs-data contains 1235 images taken from flying-cam devices (Phantom 4 Pro, Airiestronic XT8) with very high resolutions in range of 12–20 Mpx. The captured areas include of mountains, forests and flats. For the safety, the images were captured with at least 5 m from the devices with many different shooting angles. Then, they were uploaded automatically into our server in which images was labeled by used LabelIMG[2]. The dataset is composed of five components as described in the Table 1 and visualized in Fig. 2.

Table 1. Dataset description with training and testing sets.

Datasets	Numbers				
	Single insulators	Double insulators	Corona rings	Discharge guns	Insulator fault
Training set	904	868	2153	1212	213
Testing set	226	217	539	304	54

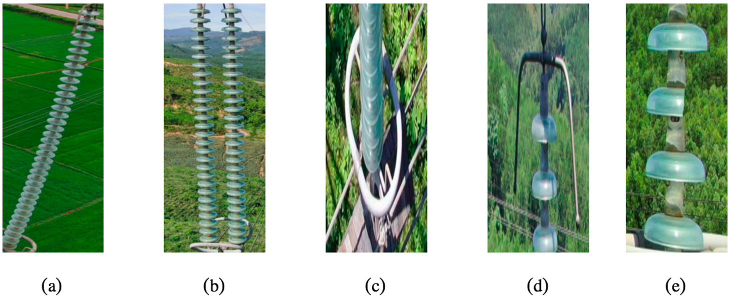

Fig. 2. Visualization of the dataset composing of: (a) Single insulator string, (b) Double insulator string, (c) Corona ring, (d) Discharge gun, (e) Insulator fault.

In our dataset, we considered the faulty parts as separate objects from original images. For example, we carefully annotated broken insulators from single/double insulator strings to create an insulator fault set with 213 and 54 images in training and testing sets, respectively. Followed this step, it is common in literature to use argumentation techniques to enrich the data.

4 Multi-object Detection

As reminder with the advent of the CNN network, object detection and recognition have been obtained a new breakthrough Fig. 3. Currently, there are many

[2] https://github.com/tzutalin/labelImg.

models using CNN to detect objects: R-CNN, Fast R-CNN, Faster R-CNN [15], YOLO [16].

We present here two different approaches to deal with our specific dataset. These approaches are well-known methods in multi-object detection in general and power transmission component inspection in detail as based on Faster R-CNN in Sect. 1, YOLOv4 in Sect. 4.2.

method	# box	data	mAP	areo	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv
SS	2000	07	66.9	74.5	78.3	69.2	53.2	36.6	77.3	78.2	82.0	40.7	72.7	67.9	79.6	79.2	73.0	69.0	30.1	65.4	70.2	75.8	65.8
SS	2000	07+12	70.0	77.0	78.1	69.3	59.4	38.3	81.6	78.6	86.7	42.8	78.8	68.9	84.7	82.0	76.6	69.9	31.8	70.1	74.8	80.4	70.4
RPN*	300	07	68.5	74.1	77.2	67.7	53.9	51.0	75.1	79.2	78.9	50.7	78.0	61.1	79.1	81.9	72.2	75.9	37.2	71.4	62.5	77.4	66.4
RPN	300	07	69.9	70.0	80.6	70.1	57.3	49.9	78.2	80.4	82.0	52.2	75.3	67.2	80.3	79.8	75.0	76.3	39.1	68.3	67.3	81.1	67.6
RPN	300	07+12	73.2	76.5	79.0	70.9	65.5	52.1	83.1	84.7	86.4	52.0	81.9	65.7	84.8	84.6	77.5	76.7	38.8	73.6	73.9	83.0	72.6
RPN	300	COCO+07+12	**78.8**	**84.3**	**82.0**	**77.7**	**68.9**	**65.7**	**88.1**	**88.4**	**88.9**	**63.6**	**86.3**	**70.8**	**85.9**	**87.6**	**80.1**	**82.3**	**53.6**	**80.4**	**75.8**	**86.6**	**78.9**

Fig. 3. Results on PASCAL VOC 2007 test set with Fast R-CNN and VGG-16.

4.1 Faster R-CNN-based System

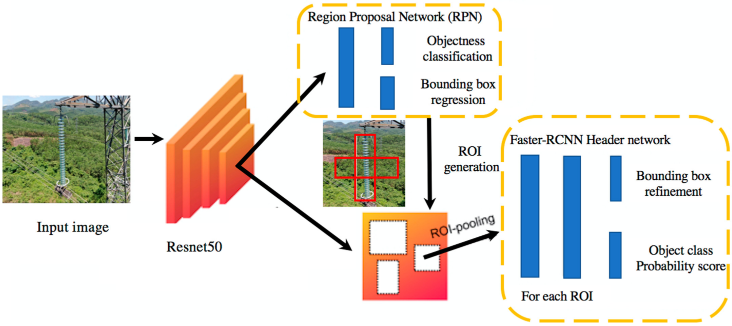

Fig. 4. The detection process bases on Faster R-CNN model.

With the advantage of the Faster R-CNN model [15], we have adopted it here of detection and recognition process with power transmission system devices (insulators, corona rings, discharge guns) Fig. 4. There are several works on this model [3, 10, 13]. We employ the Resnet50 network to extract image features. We use this network based on advantages of being not too large in size and providing high accuracy Table 2.

4.2 YOLO-based System

Recently, YOLOv4 [17] has appeared as the state-of-the-art object detector. For this reason, we employ the pre-trained weights of YOLOv4 model for our dataset.

Table 2. Pre-trained model (available online https://keras.io/api/applications/).

Model	Size	Top-1 accuracy	Top-1 accuracy	Parameters	Depth
Xception	88 MB	0.79 %	0.945%	22,910,480	126
VGG16	528 MB	0.713 %	0.901%	138,357,544	23
ResNet50	98 MB	0.74 %	0.92%	25,636,712	26
ResNet101	171 MB	0.764 %	0.928%	44,707,176	26
InceptionV3	92 MB	0.779%	0.937 %	23,851,784	159

For convenience, we remind here the main points of this model. YOLOv4 includes of three key parts as: backbone - CSPDarnet53, neck - SPP (Spatial pyramid pooling), and head - YOLOv3. YOLOv4 uses BoF (Bag of Freebies), BoS (Bag of Specials) for backbone and detector. This approach is shown in Fig. 5.

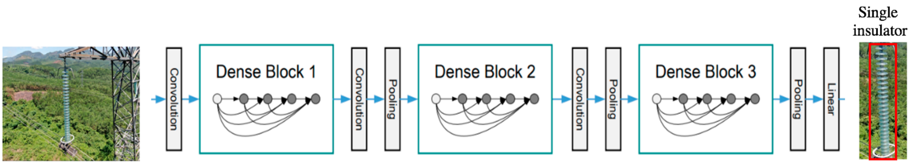

Fig. 5. The detection process based on YOLOv4 model [17].

5 Our Proposed System

Our overall system is provided in Fig. 6. Image acquisition is the first step of the inspection process in which captured images of power transmission components using a camera on the UAV. All captured images used in the paper were taken

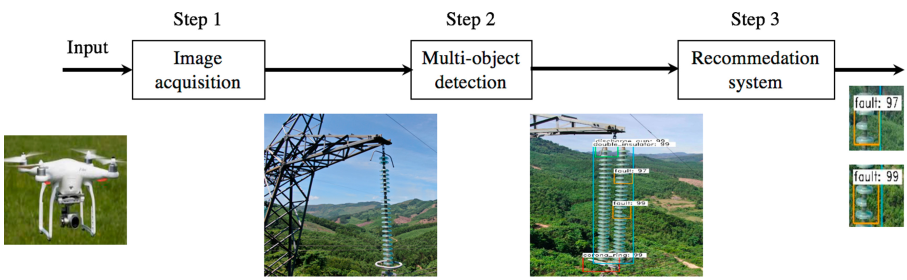

Fig. 6. Our proposed system for power transmission inspection based on CNN with three main steps.

on the Vietnam Electricity of The National Power Transmission Corporation (EVNNPT). The captured areas include of mountains, forests, making it difficult to traditional approaches. The collected images are then sent to our sever to trigger the second step.

In the second step, we propose to apply YOLOv4-based system. Compare to the Faster R-CNN-based system, YOLOv4-based system is faster and more accurate in average [17]. It therefore can perform this system in a near real-time processing. The acquired images are passed through this system to detect five given objects as listed in Table 1. In the third step, the situation of each component is observed by cropped it from detection regions. In the scope of this paper, we only used a case study for informing the broken insulators.

It is worth noting here that most of the state-of-the-art works on detecting faulty devices processed with segmentation problems [5,6,10]. However, the approaches are quite sensitive with complex backgrounds having a large similarity to the neighboring pixels. Therefore, we take into account the broken parts as particular objects (see Sect. 1). When the broken part is detected, this region will be located and informed. A report will guide for power companies to find exactly locations of the insulator faults.

6 Experiments

In this section, we present the performance evaluation of our system on the proposed dataset. For the multi-object detection task, we only concentrate on the step 2 of our system Fig. 6. Sections 6.1 and 6.2 detail the experimental setups and metrics. At last, Sect. 6.3 presents the results.

6.1 Experimental Setups

At the training step, the pre-trained weights of YOLOv4 was cloned from Darknet. By experimental loops and recommendation for customizing parameters from original paper with YOLOv4 [17], we modified the config file as follows: batch = 64, subdivisions = 32, max-batches = 10000, step = 8000,9000, classes = 5, filters = 30. In order to improve the robustness of the training model, image augmentation was deployed as hue variation (in range 0.9 to 1.1), exposure variation (in range 1 to 1.5), and rotation (in range $-60°$ to $60°$). In addition, we developed in dilation and erosion of training samples to enrich the dataset. The model was obviously trained by an open-source python environment using Darknet library.

6.2 Metrics

Object detection systems predict the outcome as bounding boxes and labels. For each bounding box we measure the overlap between the predicted box with the ground truth bounding box called the Intersection over Union (IoU) [18]. For each class we calculate precision and recall based on the IoU threshold (0.5). We define a prediction to be a True Positive (TP) if the IoU > 0.5, False Positives (TP) if the IoU > 0.5 and False Negative (FN) if a ground truth not detected. Average precision (AP) is the measurement calculated based on the Precision (P) - Recall (R) curves for all data points (M).

$$P = \frac{TP}{TP + FP} \qquad R = \frac{TP}{TP + FN} \qquad AP = \sum_{m=1}^{M} (R_m - R_{m-1}) P_m$$

Table 3. Results on three approaches on the detection task (shown in the step 2 of Fig. 6), the proposed system using YOLOv4 outperforms other approaches on our dataset with big gaps.

Data	Objects				
	Fault insulator	Single insulator	Double insulator	Corona ring	Discharge gun
VGG16-based system- Average precision (AP %)					
Training set	71.8 %	53.5%	64.6%	79.3%	84/37%
Testing set	53%	61.2%	70.5%	61.65%	50.7%
Faster R-CNN-based system - Average precision (AP %)					
Training set	82.8%	80.8%	89.9%	68.6%	57.6%
Testing set	71%	64.3%	71.4%	50.5%	49.1%
Proposed system - Average precision (AP %)					
Training set	99.95%	99.54%	99.8%	99.66%	99.42%
Testing set	88.54%	93.25%	93.32%	81.06%	88.51%

6.3 Results

Our implementations were test on ThinkLabs-data with Tesla K80 GPU, 12 GB RAM architecture. The results are given in Table 3. It can see that our approach outperforms others on all objects. The results on single/double insulators are performed in best about 93.3%. It is true that the object detection system could fail to detect single insulator, double insulator, fault insulator in different scenarios. It could result from the different shooting angles from the flying-cam device. Next, the complexity of the background image also causes incorrect prediction results. Compare to the performance on single/double insulators, result on fault insulator is a bit smaller with 4% of gap. It could be explained that the faulty parts in the insulator strings are quite smaller, they could disappear in some angles of the flying-cam devices. Corona rings and discharge gun also occur several errors. This could result from the same reasons as the faulty parts of the insulators. The successful detection is visualized in Fig. 7.

In addition, Faster RCNN-based approach provides regions of interest for doing convolution, it is strongly dependent on the sizes of anchor points. Meanwhile, YOLO-based approach does detection and classification in one shot, it could affect the final results on our dataset.

For clarity, we measure the processing time in the testing phase. The results are shown in Table 4. It can be seen that our approach is almost ten times as fast as the VGG16-based system and five times as fast as the Faster R-CNN system. It could perform as a near real-time processing.

Table 4. Results on the processing time of three different approaches for the detection tasks (in second).

Methods	VGG16-based system	Faster R-CNN-based system	Our system
Processing time	1.12 (s)	0.68 (s)	0.158 (s)

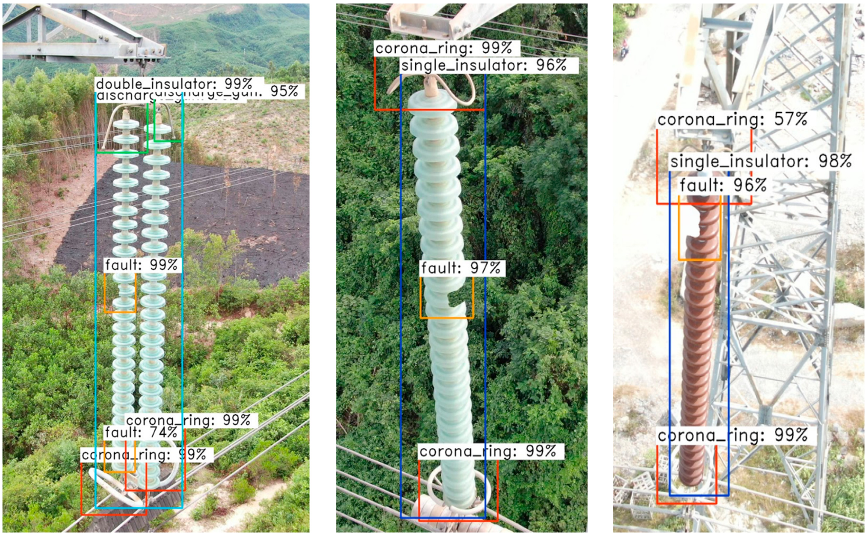

Fig. 7. Visualization of our detection results.

7 Conclusions and Perspectives

In this paper we proposed a new standard dataset of electric power equipment for high-voltage transmission line inspection. Then, different approaches are re-implemented to automatically detect and recognize objects with a special focus on the insulator fault detection. Our proposed system performed as the best candidate for this task.

The proposed system performs in a near real-time processing. It could be embedded directly into UAV for detection and recognition tasks. Then, a new method for scheduling the flying direction could be embedded into the UAV to improve the quality of detection and recognition tasks.

Acknowledgments. This research was funded by ThinkLABs R&D[3] and the sample data was supported by EVN NPT2[4] in Da Nang.

References

1. Pagnano, A., Höpf, M., Teti, R.: A roadmap for automated power line inspection. Maintenance and repair. Proc. Cirp **12**, 234–239 (2013)
2. Shakhatreh, H., Sawalmeh, A.H., Al-Fuqaha, A., Dou, Z.: Unmanned aerial vehicles (UAVs): a survey on civil applications and key research challenges. IEEE Access **7**, 48572–48634 (2019)
3. Liu, X., Miao, X., Jiang, H., Chen, J.: Review of data analysis in vision inspection of power lines with an in-depth discussion of deep learning technology. Preprint arXiv:2003.09802 (2020)
4. Zhai, Y., Chen, R., Yang, Q., Li, X., Zhao, Z.: Insulator fault detection based on spatial morphological features of aerial images. IEEE Access **6**, 35316–35326 (2018)
5. Wang, Y.L., Yan, B.: Vision based detection and location for cracked insulator. Comput. Eng. Design **35**(2), 583–587 (2014)
6. Oberweger, M., Wendel, A., Bischof, H.: Visual recognition and fault detection for power line insulators. In: Computer Vision Winter Workshop (2014)
7. Hao-ran, J., Lin-jun, J.I.N.: Recognition and fault diagnosis of insulator string in aerial images. J. Mech. Electr. Eng. **32**(2) (2015)
8. Yin, J., Lu, Y., Gong, Z., Jiang, Y., Yao, J.: Edge detection of high-voltage porcelain insulators in infrared image using dual parity morphological gradients. IEEE Access **7**, 32728–32734 (2019)
9. Wang, M., Du, Y., Zhang, Z.R.: Study on power transmission lines inspection using unmanned aerial vehicle and image recognition of insulator defect. J. Electron. Meas. Instr. **26**, 1862–1869 (2015)
10. Han, J., et al.: A method of insulator faults detection in aerial images for high-voltage transmission lines inspection. Appl. Sci. **9**(10), 2009 (2019)
11. Siddiqui, Z.A., Park, U.: A drone based transmission line components inspection system with deep learning technique. Energies **13**(13), 3348 (2020)
12. Kang, G., Gao, S., Yu, L., Zhang, D.: Deep architecture for high-speed railway insulator surface defect detection: denoising autoencoder with multitask learning. IEEE Trans. Instr. Meas **68**(8), 2679–2690 (2018)
13. Ma, L., Xu, C., Zuo, G., Bo, B., Tao, F.: Detection method of insulator based on faster R-CNN. In: Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), pp. 1410–1414 (2017)

[3] https://thinklabs.vn/vn/.
[4] https://www.npt.com.vn.

14. Tao, X., Zhang, D., Wang, Z., Liu, X.: Detection of power line insulator defects using aerial images analyzed with convolutional neural networks. IEEE Trans. Syst. Man Cybern. Syst. 1–13 (2018)
15. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems, pp. 91–99 (2015)
16. J., Redmon, S., Divvala, R., Girshick: You only look once: unified, real-time object detection. In: CVPR, pp. 779–788 (2016)
17. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: YOLOv4: optimal speed and accuracy of object detection. Preprint arXiv:2004.10934 (2020)
18. Nguyen Dinh, C., Delalandre, M., Conte, D., Pham, T.A.: Fast RT-LoG operator for scene text detection. J. Real-Time Image Process. **18**(1), 19–36 (2020). https://doi.org/10.1007/s11554-020-00942-7