A Self-Constructed CNN Classifier for Keyhole Detection and Location

Junjie Ye
School of Electrical Engineering
Southeast University
Nanjing, China
806731385@qq.com

Bingtuan Gao
School of Electrical Engineering
Southeast University
Nanjing, China
gaobingtuan@seu.edu.cn

Hao Chen
State Grid Jiangsu Electric Power
Maintenance Branch Company
Nanjing, China
pingfengma@126.com

Weilun Xu
School of Electrical Engineering
Southeast University
Nanjing, China
220182738@seu.edu.cn

Linlin Zhong
School of Electrical Engineering
Southeast University
Nanjing, China
linlin@seu.edu.cn

Chuande Liu
School of Electrical Engineering
Southeast University
Nanjing, China
925683993@qq.com

***Abstract*—With the development of the demand for robots to have the ability of opening and unlocking, a real-time detection algorithm for keyhole is proposed in this paper, which combines a fast circle detection based on circle edge feature and a keyhole classifier using self-constructed CNN. The algorithm needs three steps: CNN construction, circle detection and keyhole classification of local circle image. Firstly, collecting keyhole images as positive samples, collecting non-keyhole images and random images as negative samples. Then, constructing a CNN classifier to solve the classification problems and training it based on the samples collected before. Consequently, detecting the real-time collected image whether it has circles and where they are. Furthermore, extract images might have circle area and using the CNN keyhole classifier trained before to determine whether each of them is a keyhole image or not. The experimental results show that using the same training and testing samples, the accuracy of self-constructed CNN classifier(>98 %) is about 1 % lower than that the CNN classifier using AlexNet migration learning network (>99 %), but its time use is about 95% less than the AlexNet and its net size is about 99.5% less than the AlexNet.**

Keywords—keyhole detection; Hough circle detection method; CNN; machine vision

2020 IEEE Sustainable Power and Energy Conference (iSPEC) | 978-1-7281-9164-5/20/$31.00 ©2020 IEEE | DOI: 10.1109/ISPEC50848.2020.9351106

I. INTRODUCTION

With the development of modern control technology, robots begin to replace human in repetitive, complex and high-risk task. However, in the scene of factory inspection, indoor emergency repair, home service, etc., robots often need to realize the operation of opening and unlocking independently; thus, the demand for robots with the ability of opening and unlocking is constantly increasing [1-2]. For example, the inspection robots of substation have been widely used now, but they can only carry out the basic external monitoring work for the instrument [3]. Their insufficiency of operation ability for the on-site equipment of substation makes them difficult to monitor the internal equipment status of the power box and cabinet at the substation site, leading to the lack of ability to monitor the power components and to handle the accident. Currently, detection and other work still need people, do not reduce the labor cost, and people maybe injured when the substation is in danger. Therefore, the demand for door-opening and unlocking robots is constantly increasing [4]. This paper studies the most basic problem: keyhole identification and location.

There are few existing literatures on keyhole recognition and location technology. The research contents of reference [4] and [5] are close to this keyhole detection. In reference [4], HOG-SVM (Histogram of Oriented Gradient- Support Vector Machine) is used to detect the fusion features of door lock and door handle to solve the door-opening question. However, in reality, door handle and door lock may not exist on the same door at the same time, and the types of keyhole training materials in the reference are single. Reference [5] studies the location of container keyhole, which has small application scope and single keyhole type. Moreover, it studies the keyhole detection method based on traditional pattern recognition and YOLO (You Only Look Once) neural network. The robustness of traditional pattern recognition is relatively poor and the end-to-end YOLO detection algorithm has a large amount of computation. Therefore, this paper proposes a machine vision-based keyhole detection algorithm under uncertain, which is applied to the case when robot needs to unlock independently in the unknown scene.

II. PROPOSED METHOD

A. Modeling and Simplification

Considering the scene features of robot unlocking in unknown place, this paper considers that the keyhole recognition and location algorithm should have the following features:

- Accuracy. The robot needs to locate the specific position of the keyhole relative to the robot accurately before unlocking.
- Rapidity. The time cost of detecting the keyhole should not be too long, so as not to affect the overall response speed of the system.
- Generality. The scene for robot unlocking is unknown, and the type of keyhole is unknown, so it should have generalization ability for keyhole detection.

The task of opening and unlocking the door by robots has a series of operations including finding the door, walking to the front of the door, identifying and locating the keyhole, and controlling its mechanical arm to unlock. Therefore, before solving the problem of keyhole recognition and positioning, this paper makes the following three assumptions for the robot unlocking:

- Assumption 1: Considering the characteristics of keyhole structure, this paper holds that there must be a circle outside the keyhole.
- Assumption 2: In this paper, all the keyholes studied are installed in vertical plane (such as door lock, closet lock, etc.).
- Assumption 3: In this paper, we default that robot has walked to the front of the vertical plane which having keyhole and is facing to the vertical plane (i.e. facing to the keyhole) before it starts to locate the keyhole.

According to the assumption mentioned above, we can set that the scene of the robot before unlocking as Fig.1, which has the following scene features:

- The robot is facing the keyhole, so the attitude of the keyhole does not need to be considered.
- The distance between the robot and the vertical plane should not be more than 2 meters to ensure that the keyhole can be identified effectively and the mechanical arm can be controlled for unlocking operation.

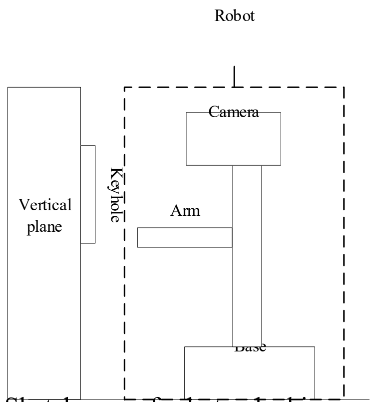


Fig. 1 Sketch map of robot unlocking scene

Thus, according to a series of assumptions and scene features mentioned above, we divide the keyhole detection task into two steps: circle detection and keyhole identification. Dividing the raw task into two smaller tasks can reduce the amount of computation, which is important to the rapidity of the algorithm. Therefore, we will solve these two tasks in *B* and *C* parts.

B. Circle detection

In order to reduce the amount of computation of images and prepare for the boundary labeling, it is necessary to preprocess the image before circle detection. The image preprocessing of this algorithm includes grayscale and local threshold binarization.

The RGB three-channel color image obtained by the camera is multiplied by different weights and becomes grayscale image (weighted average method). Then, the range of pixel value is 0-255. In order to reduce the amount of computation, the grayscale image need to be converted into black-and-white image. The simplest method is global threshold binarization. First, the most appropriate segmentation threshold value is calculated according to OTSU [6], and then the binarization is completed as:

$$B_{ij} = \begin{cases} 0, G_{ij} < q \\ 1, G_{ij} > q \end{cases} \tag{1}$$

where q represents the selected gray threshold, G_{ij} represents the gray value of a pixel on the image, and B_{ij} represents the binary pixel value of the corresponding pixel.

The global threshold binarization is not suit for selecting an appropriate segmentation threshold for the image with partial light and partial dark areas (such as the image with uneven light reception). In this paper, we use Niblack method [7] to calculate the segmentation threshold q according to the mean value and variance of the pixel gray level in the $R\times R$ region around the pixel, and then the image can be binarized.

According to Assumption 1 and reflective characteristics of keyhole metal material, we determine that the area contains keyhole is white after binarization. Therefore, we use Two Pass method [8] to calibrate the white connected area of the binarization image. Then, the boundary labeling is completed by recording the boundary coordinate set and the label of the connected domain.

After the boundary labeling, we need to filter out the circle area Therefore, we build the screening conditions in combination according to the scene features and the geometric features of the circle. We have constructed two quick screening conditions.

There are a large number of connected domains in an image. We build simple screening conditions to filter out most of the connected domains that do not meet the requirements.

Calculate the area of the connected domain from recorded boundary coordinate assemble and the connected domain with too small area has no meaning to be detected, i.e. the connected domain that does not meet the requirements of the following formula should be filtered out.

$$\lambda = \frac{A(i)}{A_{\text{sum}}} > 0.001 \tag{2}$$

where $A(i)$ represents the pixel area of the ith connected domain, and A_{sum} represents the total pixel area of the image.

Then, according to the geometric characteristics of circle that the length of diameter in any direction of the circle should be same, a screening condition is constructed to filter out the connected regions whose span difference in the x and y direction is too large, i.e. remove the connected domain that do not meet the following formula:

$$span = x_{\max} + y_{\min} - x_{\min} - y_{\max} \tag{3}$$

where $x_{\max}$ and $x_{\min}$ represent the maximum and minimum values in the x direction of the boundary coordinate matrix, and $y_{\max}$ and $y_{\min}$ represent the maximum and minimum values in the y direction of the boundary coordinate matrix, *span* means the difference of boundary coordinate in x and y direction.

For the diameters of different circles are different, we set up to normalize the span in x and y direction:

$$\mu = \frac{span}{x_{\min} + x_{\max} + y_{\min} + y_{\max}} < 0.01 \quad (4)$$

where μ is the normalized difference of boundary in x and y direction.

We assume that if μ>0.01, the boundary difference is so large that it is impossible that the boundary is similar to a circle.

C. Keyhole detection

After detecting the circle area of the image, we extract the partial image contains circle area, then we need to identify whether the image is a keyhole or not. Considering the variety and color of the keyhole in the environment we are targeting, the identification method based on keyhole geometric features is inevitable to be unable to deal with the situation with high uncertainties. Therefore, we choose CNN (convolution neural network) to realize the identification of keyhole images.

Different from the traditional neural network model, CNN is specially proposed for solving image classification problems. It convolutes the local image through convolution kernel, so as to effectively extract the image local features to achieve image classification. CNN's workflow is shown in the Fig.2.

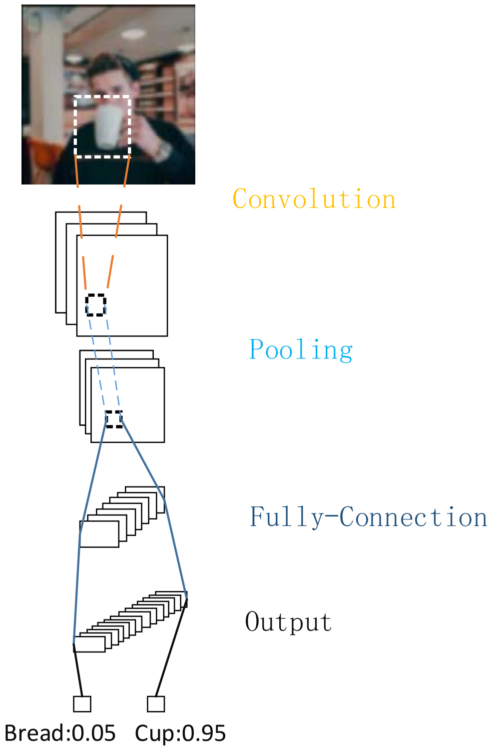


Fig. 2 CNN's classification principle

CNN is usually used in image classification to train the keyhole classifier [9]. The most popular neural network model is AlexNet [10] network, which can classify 1000 kinds of images with higher accuracy than the naked eye. Because of the mature network structure and excellent recognition effect of AlexNet, many people will carry out migration learning based on it. They will rebuild their own network by replacing the layers since full connection layer to their new layers and train them according to the research problems.

- CNN's preparation and training

In the problems studied in this paper, algorithm needs to realize image classification quickly, which is different from the AlexNet. Therefore, a new convolution neural network is built and trained in this paper. Then main steps are shown in Fig.3.

First, since there is no existing training set of keyhole samples in the world, we need to build our own training set for convolutional neural network. We collect a variety of keyhole area images, make appropriate morphological transformation and get 1040 keyhole images, which are uniformly scaled to the size of 227×227×3, as the positive sample set of keyholes. After that, we collect 1089 pictures including non-keyhole pictures including pictures having circle, random pictures and geometric pictures, which are also scaled to 227×227×3, as the negative sample set of keyholes. As shown in Fig.3, the completed training set of keyholes is constructed. Then we label each picture mentioned before about whether it belongs to keyhole or no, and randomly select 70% as the training set, and the remained 30 % images is the test set.

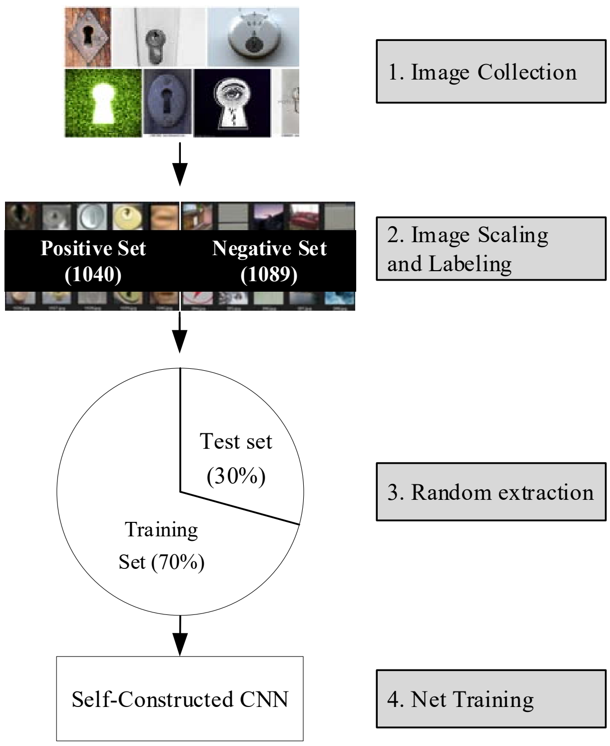


Fig. 3 CNN's preparation and training

- CNN’s network structure

According to the format of training set pictures and CNN’s network structure features, we build our own convolutional neural network. The size of input image is the same as the AlexNet, thus CNN’s network structure. We build a CNN with three convolution layers, three maximum pooling layers, four activation layers and some other layers. The specific network parameters of CNN are shown in TABLE 1 and the specific process of image processing is shown in Fig.4.

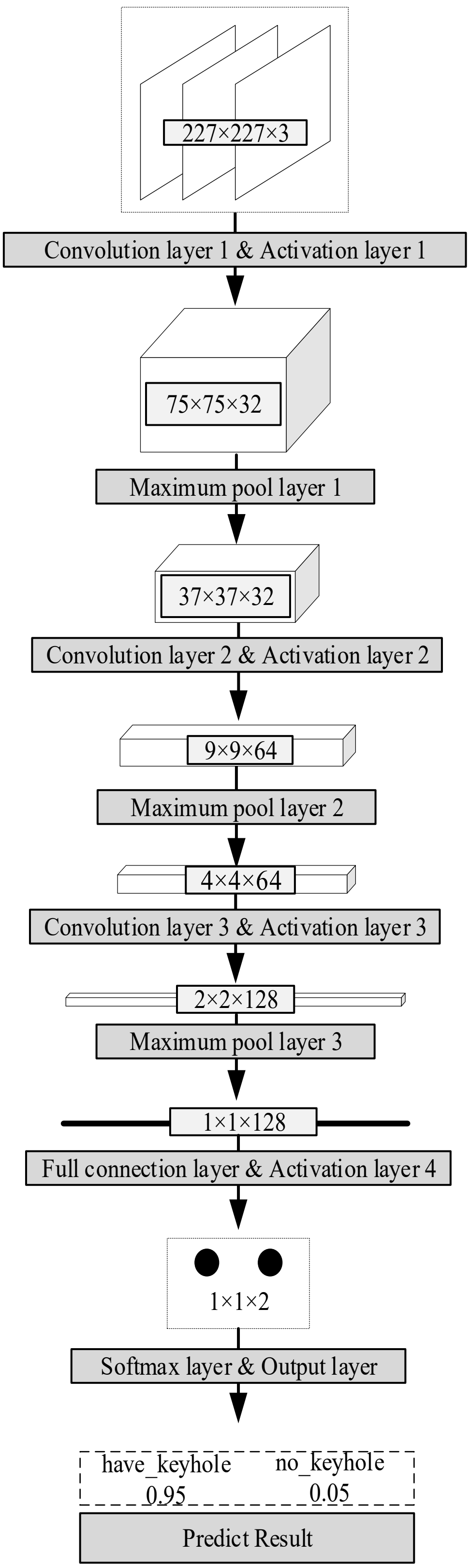


Fig. 4 Self-constructed CNN's network structure

III. EXPERIMENT

In order to test the effectiveness, advantages and disadvantages of the circle detection and keyhole identification algorithm and verify their feasibility in this paper, we build a simulation platform on the computer, and compare it with other algorithms having same or similar functions. Our experimental platform is MATLAB R2017b installed in the windows 10 system running on Inspiron 3559 notebook computer of Dell. The computer is configured as Intel i5-6200u CPU, with the main frequency of 2.30 GHz and 8.00 GB RAM.

In the control experiment to test the effect of classifier in keyhole identification task, we choose the migration learning neural network based on AlexNet. For the convenience of writing, the convolution neural network built by ourselves will be called "SelfNet" in the following text, and the convolution neural network based on AlexNet through migration learning will be called "TransNet". Both SelfNet and TransNet are trained using the sample set mentioned before. For the train set and test set need to be divided randomly, this paper has trained many nets and choose the best network to save. When we test these two nets, we call the network to test all the pictures and the test results are shown in TABLE.2 and Fig.5.

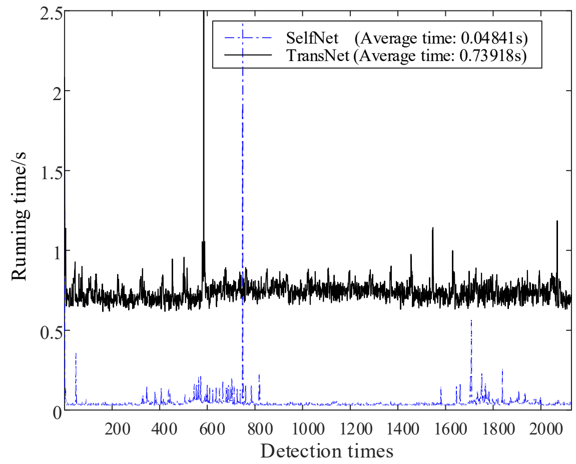


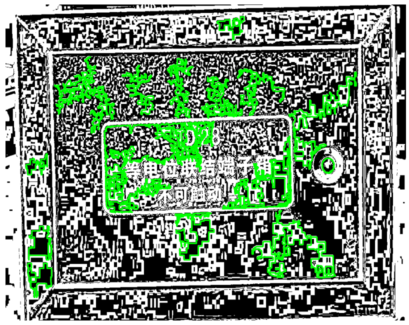
Fig. 5 Running time comparison

We can see that both the recognition accuracy of SelfNet and TransNet are very high, though SelfNet is about 1% lower than TransNet. However, its running time and network size are far smaller than TransNet, which shows that our self-constructed network has fast running speed and very little storage occupation while ensuring the accuracy, which can reduce the hardware requirements for the robot to run this network, and can better adapt to the robot to carry out real-time detection of keyhole.

Finally, we select an image to conduct a complete keyhole detection, then the keyhole detection of visible image is carried out on the MATLAB platform mentioned before, and the results are shown in Fig.6.

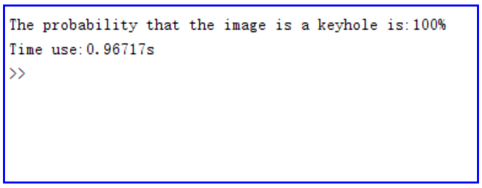
The subgraph (a) is the original image we selected. After preprocessing, we set up the screening condition and then we get the possible local circle images which are shown as the green line in the subgraph (b). At last, we call the classification network trained before to complete the detection of the circle area about whether it is keyhole or not. The highest score area is considered to be the area where the keyhole is located, which is shown in subgraph (c). The running time and the identification score are listed in subgraph (d).

(a) Original image



(b) Circle area detection

(c) Keyhole detection



(d) Running time

Fig. 6 Complete keyhole detection

TABLE 1 Self-construct CNN's network parameters

Layer name	**Position**	**Parameter setting**
Input layer	1	Input matrix: 227×227×3
Convolution layer	2	Convolution kernel size: 5×5; Stride: 3; Channel number: 3; Convolution kernel number: 32.
	5	Convolution kernel size: 5×5; Stride: 4; Channel number: 32; Convolution kernel number: 64.
	8	Convolution kernel size: 5×5; Stride: 1; Channel number: 64;Convolution kernel number: 128.
Activate layer	3、6、9、12	/
Maximum pool layer	4	Pool size: 3×3; Stride:2
	7	Pool size: 3×3; Stride:2
	10	Pool size: 2×2; Stride:2
Full connection layer	11	Size: 128
	13	Size: 2
Softmax layer	14	/
Output layer	15	Category: {“have_keyhole”,”no_keyhole”} Output vector size: 2×1

TABLE. 2 Network performance comparison

	Identification results	**Keyhole**	**Not keyhole**	**Accuracy**	**Average time use**	**Network size**
SelfNet	Classified as keyhole	1030	16	98.78 %	0.045 s	999 Kb
	Not classified as keyhole	10	1073			
TransNet	Classified as keyhole	1040	4	99.81 %	0.836 s	202 Mb
	Not classified as keyhole	0	1085			

IV. CONCLUSION

In this paper, we mainly study the keyhole recognition and positioning technology based on machine vision, which can be used to solve the problem of opening and unlocking which need to be considered in substation robots. First of all, we simplify the problem and make some appropriate assumptions. Then, according to the actual needs, we divide the keyhole detection task into two steps: circle detection and keyhole identification. As for circle detection, this paper mainly constructs the screening conditions according to the circle features, and extracts the connected regions that meet the requirements from visible images which was processed before. As for the keyhole identification problem, this paper refers to the existing results and relative works and then builds a simple CNN by ourselves. After training with the same samples, the prediction accuracy of the CNN proposed in our paper is comparable to the result using transfer learning network based on AlexNet. But the network size and detection speed of the CNN described in this paper are far less than the latter.

Finally, this paper selects an appropriate visible image and runs the complete algorithm on the MATLAB platform to realize the keyhole detection of the visible image, which verifies the feasibility and effectiveness of the algorithm in the low configuration hardware environment.

REFERENCES

[1] C. Y. Ma, H. B. Gao, L. Ding, J. G. Tao, K. R. Xia, H. T. Yu, “Optimal energy consumption for mobile manipulators executing door-opening task,” Mathematical Problems in Engineering, vol. 2018, 2018.

[2] Y. Karayiannidis, C. Smith, F. E. V. Barrientos, P. Ogren, D. Kragic, “An adaptive control approach for opening doors and drawers under uncertainties,” IEEE Transactions on Robotics, vol. 32, no.1, pp. 161-175, February 2016.

[3] P. Qian, Y. Zhang, J. M. Xu, “Research on terminal visual inspection technology of smart grid inspection robot,” Modern Electronics Technique, vol. 41, no. 18, pp. 113-116, 2018.

[4] X. Xu, “Visual localization and end effector design of mobile robot unlocking operation”. Harbin Institute of Technology, 2017.

[5] Y. D. Zhang, “Research on recognition and location technology of container keyhole based on machine vision,” Jiangsu University, 2019.

[6] H. Lamphar, “Spatio-temporal association of light pollution and urban sprawl using remote sensing imagery and GIS: A simple method based in OTSU's algorithm,” Journal of Quantitative Spectroscopy and Radiative Transfer, vol. 253, pp. 107068, August 2020.

[7] A. Chandra, S. Mondal, “Amalgamation of iterative double automated thresholding and morphological filtering: a new proposition in the early detection of cerebral aneurysm,” Multimedia Tools and Applications, vol. 76, no.22, pp. 23957–23979, 2017.

[8] X. Zhao, “Speed-up of connected component labeling algorithm and its application in paper defect detection,” Shanxi University of Science and Technology, 2019.

[9] S. Din, A. Paul, A. Ahmad, “Smart embedded system based on demosaicking for enhancement of surveillance systems,” Computers and Electrical Engineering, vol. 86, pp. 106731, September 2020.

[10] K. M. Hosny, M. A. Kassem, M. M. Fouad, “Classification of skin lesions into seven classes using transfer learning with AlexNet,” Journal of digital imaging, June 2020.