Pedestrian Detection and Type Recognition around Electric Communication Equipment

Guangshuo Chen*
Beijing Jiaotong University, Beijing, China
chenguangshuo@foxmail.com

Abstract

Today, with the increasing development of electrification, there are more and more peripheral devices that need to be monitored, such as electric communication equipment. If there are pedestrians entering the surrounding area of the monitoring area in violation of regulations, it is likely to cause equipment failure or even catastrophic personal safety accidents. Therefore, conducting research on pedestrian detection around electric communication equipment and identification of specific types of workers is of great significance for improving the economic and social benefits of enterprises. The pedestrian detection and specific worker type recognition method implemented by this design can complete the functions of reading video streams, detecting dynamic objects, judging pedestrian targets, and recognizing personnel types. It has the advantages of high accuracy, convenient configuration, low hardware requirements, and good real-time performance. This method basically meets the actual environmental requirements of this article.

Keywords: pedestrian detection, type recognition, electric communication equipment

1 Introduction

With the development of Internet of Things [18] technology, the surrounding sites with potential safety hazards such as power communication equipment deployed are a common environment. This type of environment has potentially huge dangers such as electric shock or other accidents of pedestrians who stray into the environment, or damage to equipment [24]. Ordinary pedestrians, specific workers, and surrounding site conditions need to be monitored to prevent accidents. In this type of environment, since there are fewer equipment emergencies, the personnel in the environment have become the main influencing factors. Therefore, it is necessary to monitor and identify persons entering the video area to ensure safety. More traditional monitoring and management measures need to arrange staff to patrol the area of the site. But with the Coronavirus pandemic [16], it is not appropriate to deploy too many staff. There are also methods of setting interception lines and warning bands. These methods will cause greater negligence and cause pedestrians to enter the site by mistake.

The research situation of pedestrian detection can be roughly divided into two types: one method is based on background modeling [7]. The general process is to first segment the foreground information of the image, then extract the target in the foreground information and perform further feature extraction, and then use some classifier algorithms to classify the feature information. However, in the case of chromatic aberration caused by changes in the illumination, and in the case of dense objects, and in the case of fluctuations and changes in the picture, the detection effect will be worse. Another way is based on statistical learning [22]. The features extracted for training usually contain information such as the color and pattern shape of the object, and then these extracted features are used to train a classifier

Research Briefs on Information & Communication Technology Evolution (ReBICTE), Vol. 7, Article No. 9 (October 10, 2021)
DOI:*10.22667/ReBiCTE.2021.10.10.009*

*Corresponding author: Beijing Jiaotong University, No. 3, Shangyuan village, Haidian District, Beijing, China, Tel: +86-17733652726

suitable for pedestrian detection. In 2005, Dalal, a researcher in related fields, proposed a pedestrian detection algorithm based on HOG features and support vector machines, which is a representative of the second method [12].

At present, there are also many methods used in the field of target type recognition. For the pedestrian type recognition in this article, it can be recognized based on the pedestrian's clothes, the location area, or the environmental conditions around the pedestrian. In recent years, there are also abundant research theories for judging the color characteristics of images. For example, in 2010, Farenzena [2] et al. used the color information of the image as the main recognition feature to classify images. In this article, the staff will wear special color clothes, so you can also consider using the image color. characteristic. Bazzani et al. also focused on the color features of the image, but the method is characterized by obtaining multiple images and then comparing them, combining multiple image color features for comparison. The accuracy of the recognition is compared with the previous ones. The accuracy and efficiency of the method have been improved.

The research content of this topic plans to realize the transmission of video data in the surrounding environment of electric communication equipment, pedestrian detection and identification of specific types of workers.

2 Dynamic Object Detection

Most organisms living in nature require the participation of the visual system. Most of the information acquired is also collected by the visual system, and the visual attention mechanism can quickly capture the "interested" area when faced with massive amounts of information [11].

Dynamic object detection is to separate the changing foreground from the static background. Comparing the current image with the image assumed to be a static background, and then distinguishing the area with obvious changes, it can be considered that there is a moving object in the area. For dynamic object detection, common methods include GMM, GMG, frame difference method, optical flow method, and background subtraction method [21]. Combined with the condition that the background area of the image in this article remains unchanged, the background subtraction method [13] is used to determine whether there is any dynamic object movement in the video stream. The main functions implemented include video stream information input, skip frame reading, image processing and dynamic object detection.

The process description of the dynamic object detection module is as follows: 1st, the module starts to run. Two methods can be adopted for data acquisition: linking to an IP camera and reading video files. Both methods can be used. For the convenience of introduction, this article uses to read the video file; 2nd, after obtaining the video stream information, because the first frame of the video may be damaged or other interfering factors, this module has set a period of time delay. It is to facilitate the selection of a suitable background image, and use the first frame image after the time delay as the background image; 3rd, as the video files recorded by the existing cameras are generally 30fps, 60fps or even higher, but for the scene in this article, it is not necessary Use such a high frequency. Therefore, this module has performed frame skipping processing on the video file, and the frequency of reading the video can be changed by changing the set value. Currently, 1 frame of image is read every 4 frames for dynamic object detection; 4th, the background image is read with the new Preprocessing of the image, the unprocessed image contains too much useless information, in the follow-up comparison will not only reduce the operating efficiency, but also affect the accuracy. Firstly, the image is scaled, then grayscaled, Gaussian blur and denoising are performed to facilitate the dynamic object detection of the picture; 5th, compare the processed background image with the current video image read, first calculate the two images The difference map finds the pixel difference, and then the difference image is changed to a binary image by

setting the threshold, and then the part that has changed is selected by the frame and the image range is judged whether the image range meets the conditions, and the excessive or small changes are removed. If there is a change that passes the abovementioned filtering conditions, it is considered that a dynamic object is generated in the current frame, and the current frame image is saved for subsequent pedestrian recognition steps.During the transmission of video data, the reliability of various network components also faces challenges [14].

2.1 Image Grayscale Algorithm

The image is composed of many pixels, and the color of each pixel is represented by the three values of R, G, and B. Therefore, a pixel has three colors: R matrix, G matrix, and B matrix. Vector matrix. A series of operations such as the grayscale processing of the image [10] is actually the operation of the pixel matrix in the image. Graying is to make each pixel meet the condition of R=G=B, and this value is the gray value of the pixel. Perform a weighted average calculation on the three components in the pixel matrix of the image to be processed with the weights set in advance.

$$Gray(i,j) = R(i,j) * 0.299 + G(i,j) * 0.587 + B(i,j) * 0.114$$

A more reasonable weighted gray value can be obtained after the weighted average operation according to the above formula.

2.2 Gaussian Blur Algorithm

Gaussian Blur [23] is a data smoothing technique, which is often used in computer vision image processing. It uses the function formula of Gaussian normal distribution. When performing image processing, the images are all two-dimensional, so a two-dimensional normal distribution function is required.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{(-(x^2+y^2)/2\sigma^2)}$$

With the two-dimensional normal distribution function, the pixel weights of each pixel can be calculated first. In order to calculate the weighted average, the sum of the weights of the pixels needs to be 1. In this way, the weight matrix of the pixel is obtained, and then the gray value of the pixel is multiplied by the weight matrix to obtain the Gaussian blurred value. All the pixels are processed accordingly, and finally a Gaussian blurred image of the grayscale image is obtained.

2.3 Image Background Subtraction Algorithm

The functions used in this part are derived from the Python-based OpenCV library. First use the absdiff() function to calculate the pixel difference value of the two frames of images, and then use the threshold() function to ignore the pixels whose changes are below the threshold and change the difference image to a binarized image according to the set threshold. Through dilate() function fills the threshold image with holes, and then uses the findContours() function box to select the changed area, and judges whether there is a moving object according to the size of the change.

3 Pedestrian Target Detection

In the research process and application of pedestrian detection algorithms, selecting a feature that can describe pedestrians well can save a lot of time and resources for subsequent classification algorithms

for classification, and can effectively reduce the difficulty of detection. Generally, the following properties will be investigated , A suitable description feature of the constituency: discrimination; similarity; independence; balance [5].

3.1 HOG Features and SVM Algorithm

Gradient histogram features are applied in the field of pedestrian target detection [5]. The local image area pixels in the direction of the gradient histogram describe a feature of the human body. This feature can describe a good edge of the human body and is not sensitive to lighting conditions. The influence of small displacement is small.Artificial intelligence is used in industry [15]. Recognition methods for pedestrian and human target detection such as neural networks and deep learning are more dependent on the size of the training data set. Only when a large number of relevant training data sets are prepared to invest in such algorithms can pedestrian detection be accurate. The rate has improved. But this is not applicable and unnecessary for the deployment of this article on relatively simple equipment, relatively empty and single application scenarios.

The HOG+SVM algorithm [4] used in this paper is the first important achievement in the pedestrian detection algorithm proposed by Navneet and Delaunay in CVPR in 2005. Pedestrian gradient histogram feature is currently the most widely used and most successful pedestrian feature description method. The HOG + SVM algorithm can be regarded as a landmark algorithm in the field of pedestrian recognition. Although some related algorithms proposed later have different names, the general idea is to imitate and develop them.

The HOG feature extraction process is described as follows: 1st, convert the input image into a gray value image; 2nd, perform the uniformization (normalization) of the color space; 3rd, gradient calculation; 4th, project the gradient in the gradient direction of the unit; 5th, standardize the block All cells in the normalization, the normalized block descriptor is called the gradient histogram feature; 6th, collect the HOG of all blocks in the detection space. The explanation of the above steps is as follows: Step 2 generally adopts the gamma correction method, whose purpose is to reduce the influence of shadows and local illumination and reduce noise interference; Step 3 is mainly to obtain the contour information of the detected pedestrian target; Step 4 Purpose It is to provide a local image coding area; step 5 normalization can increase the compression for the light, shadow, and edge in the image; step 6 is to collect the gradient histogram features of all divided blocks in the current detection window, and Combine into the final feature vector.

3.2 NMS Algorithm

The idea of non-maximum suppression (NMS algorithm) [20] is to search for local maximums and suppress non-maximum elements. In order to reduce the number of detection and frame selection of the same person information in the same image, the NMS algorithm is added to the HOG+SVM algorithm.

The implementation steps of the classic NMS are as follows: 1st, set the confidence threshold of the target box, the threshold in this article is set to 0.65; 2nd, the target box is arranged in descending order according to the confidence of the target box; 3rd, select the target box with the highest confidence in the current ranking list Add M to the output list, and delete it in the candidate box list; 4th, calculate the IoU value of M and other boxes in the candidate box list, and delete candidate boxes that are greater than the threshold; 5th, repeat the above process in a loop until the candidate box list is Empty, return the output list. The IoU value represents the intersection ratio, as shown in Figure 1, IoU = (the area of the overlapping part of the two framed regions)/(the area of the union of the two framed regions). In the application scenario of this article, when IoU is bigger than 0.75, it can be considered as a good frame selection result.

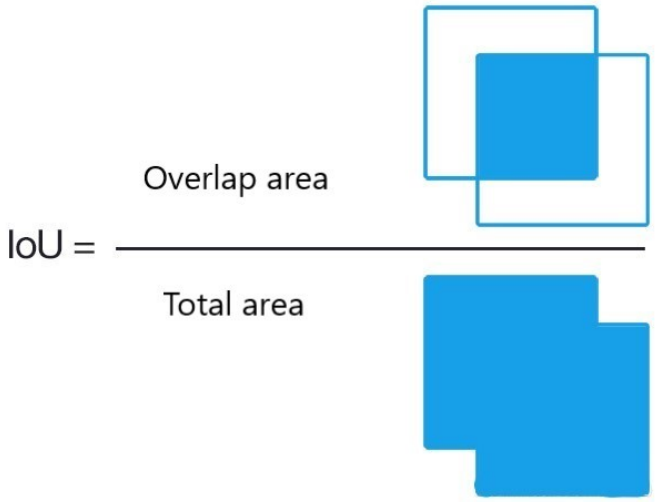

Figure 1: The meaning of IoU

3.3 Pedestrian Detection Effect

Pedestrian detection effect diagram in outdoor, indoor and night environment, as shown in Figure 2 on page 6:

4 Person Type Recognition

According to the needs of this article, the detection and type identification of people in relatively open spaces around electric equipment, combined with the relevant clothing requirements and actual application requirements of the people entering the field, design a method that uses the dichotomous K-means algorithm to first gather the color values of the people's pictures. Class algorithm, and then the target recognition method that matches the color feature of the standard person's picture, and the recognition result is obtained by matching the similarity of the image color feature [25].

4.1 Binary K-means Clustering Algorithm

Since the k center points selected by the traditional K-means algorithm are random [19], there will be a problem of converging to a local minimum. In order to overcome this problem, another clustering algorithm called bipartite K-means is proposed. The algorithm first divides all data points into one cluster, and then divides the cluster into two, and then chooses which cluster can minimize the SSE (Sum of Squared Error) value, and then Choose one of the two clusters to continue the division. The above division of clusters will be repeated until the number of clusters reaches the required value of k, which is significantly improved compared to the classification effect of the K-means algorithm [9].

The description of the process of the binary K-means clustering algorithm is shown as follows: 1st, treat all data points as a cluster; 2nd, when the number of clusters is less than k, calculate the total error for each cluster; 3rd, perform on a given cluster K-means (k=2); 4th, calculate the total error after dividing the current cluster into two; 5th, select the cluster with the smallest error for division operation; 6th, when the number of clusters is less than k, jump to step 3, loop Perform steps 3-5 until the number of divided clusters reaches the value of k.

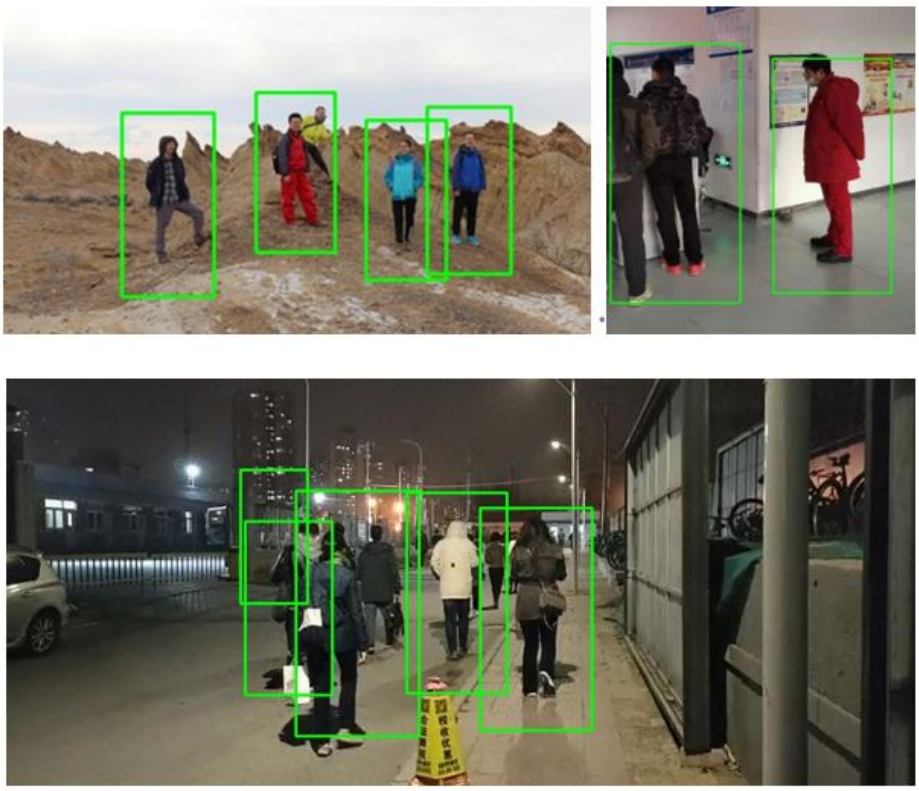

Figure 2: Effect picture

4.2 Three Histogram and Single Channel Histogram

Common image color similarity detection algorithms are as follows:

(1)The mean hash algorithm [6] calculates after processing the picture, so the calculation speed is fast and the ratio is easily affected by the size of the input image.

(2)The general steps of the perceptual hash algorithm [6] are: 1st, reduce the picture to facilitate calculation; 2nd, simplify the color; 3rd, calculate the DCT; 4th, reduce the DCT; 5th, calculate the average value; 6th, calculate the hash value.

(3)The general steps of the difference hash algorithm [6] are 1st, reduce the picture; 2nd, simplify the color; 3rd, calculate the difference value; 4th, process the difference value; 5th, merge the results to get the hash value.

The histogram has been widely used in image processing. The image color recognition algorithm in this article uses a combination of three histograms and single-channel histograms [3]. The histogram can calculate the number of pixel gray values in the image, and then calculate the three histogram of the image that needs to be compared. Compare the similarity of the histograms of two images with the single-channel histogram, and then judge the similarity of the two images. This module in this article uses the property of the histogram that only counts the number of pixel gray values, and does not need to reflect the specific location information of the pixel gray value of the image. For this demand, the histogram is just the Appropriate choice for this method.

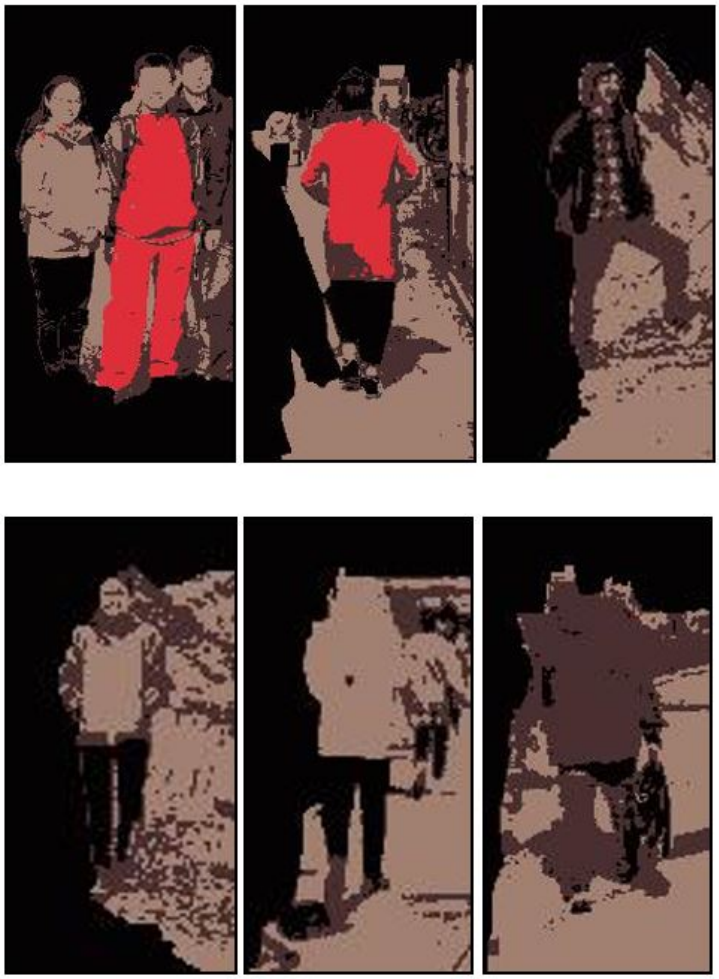

Figure 3: Effect picture after clustering

4.3 Pedestrian Color Clustering Effect

The pedestrian images extracted by the HOG+SVM algorithm are processed by the foreground extraction algorithm, and then processed by the binary K-means clustering algorithm to obtain the clustered image information, as shown in Figure 3 on page 8:

5 Simulation

For the test data set and training data set used in this article, the accuracy of model detection is above 97%. As shown in Figure 4 on page 8, the log file when training the model. The algorithm loads 3542 positive samples and 1671 negative samples when training the model. , The training model has a minimum of 151 samples for detecting errors; the accuracy of detecting pedestrians on the data set created by oneself can also reach 90% or more.

Table 1 shows the results calculated by the image color feature comparison algorithm. The explanation of the table is as follows: “1” in the prediction result indicates that the detection result is a specific worker; “0” in the prediction result indicates the detection result It is an ordinary pedestrian; “1” in the

```
2021-11-22 18:15:32,507 INFO Program runing...
2021-11-22 18:15:32,532 INFO load pos samples:3542
2021-11-22 18:15:32,533 INFO load neg samples:1671
.....
2021-11-22 21:43:01,299 INFO 47.th correction is training...
2021-11-22 21:43:01,221 INFO 46.th correction train is done,use time:321,wrong samples:151
```

Figure 4: Run log file

	Actual Result	
Predicted Result	1	0
1	9	0
0	2	17

Table 1: The test result

actual result indicates that the pedestrian is a specific worker; "0" in the actual result indicates that the pedestrian is an ordinary pedestrian.

According to the calculation formula, the accuracy rate of the image color feature comparison algorithm is 92.9%, the recall rate is 81.8%, and the F1 value is 87%.

6 Summary

The data set used in the pedestrian target detection module of the method designed in this paper is the INRIA data set, which contains a wealth of information and can cover the application scenarios of this article, the accuracy of this module on the data set is close to 97% and above, similar to the accuracy of common mainstream human body recognition algorithms. In the person type recognition module, the training and testing data sets are manually created, and the module's operating accuracy can reach up to 92%, which is higher than the accuracy of common image similarity comparison algorithms by about 10%. The data compression method is optimized to facilitate efficient data transmission between ordinary computers. This method basically meets the actual environmental requirements of this article.

Furthermore, for the problem of using camera to collect data, video data can be collected using smart unmanned aerial vehicles (UAVs) [8]. It is necessary to analyze risks in aspects such as user privacy vulnerabilities [17]. For the realization of efficient wireless bandwidth utilization to speed up data transmission, cognitive radio technology can be adopted [1].

References

[1] Y.-S. Chen, C.-H. Cho, I. You, and H.-C. Chao. A cross-layer protocol of spectrum mobility and handover in cognitive lte networks. *Simulation Modelling Practice and Theory*, 19(8):1723–1744, 2011.

[2] M. Cristani, M. Farenzena, D. Bloisi, et al. Background subtraction for automated multisensor surveillance: A comprehensive review. *EURASIP Journal on Advances in Signal Processing*, 2010(343057), August 2010.

[3] G. Dajie. *Research on Foreground Extraction and Target Detection and Tracking Algorithms in Surveillance Video*. PhD thesis, Zhejiang University, 2016.

[4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Proc. of the 2005 IEEE Computer Society Conference on Computer Vision & Pattern Recognition (CVPR'05), San Diego, CA, USA*. IEEE, June 2005.

[5] Q. Dewen. *Research on image classification algorithm and application based on support vector machine*. PhD thesis, Southwest University, 2020.

[6] H. Jiaheng, L. Xiaowei, C. Benhui, et al. Comparative study on image similarity algorithms based on hash. *Journal of Dali University*, 2(12):6, 2017.

[7] F. Kai, W. Tianhao, L. Tengfei, F. Xueying, and L. Qiang. Internal attention selection in visual working memory: Effects, characteristics and mechanisms. *Psychological Science*, 43(6):8, 2020.

[8] P. Kumar, S. Garg, A. Singh, S. Batra, N. Kumar, and I. You. Mvo-based 2-d path planning scheme for providing quality of service in uav environment. *IEEE Internet of Things Journal*, 5(3):1698–1707, 2018.

[9] J. Macqueen. Some methods for classification and analysis of multivariate observations. 5.1:281–297, 1967.

[10] G. Meihua, S. Binbin, W. Miaomiao, et al. Overview of color image graying algorithm. *Application Research of Computers*, 36(5):7, 2019.

[11] Z. Peng. *Research on target tracking based on human visual attention mechanism*. PhD thesis, Nanjing University of Aeronautics and Astronautics, 2017.

[12] Z. Ruirui. *Research and system implementation of pedestrian detection and re recognition*. PhD thesis, Anhui University, 2019.

[13] L. Shuo. *Research on moving target detection and tracking algorithm based on background subtraction*. PhD thesis, Xi'an Shiyou University, 2019.

[14] F. Song, Z. Ai, H. Zhang, I. You, and S. Li. Smart collaborative balancing for dependable network components in cyber-physical systems. *IEEE Transactions on Industrial Informatics*, 17(10):6916–6924, 2021.

[15] F. Song, Z. Ai, Y. Zhou, I. You, K.-K. R. Choo, and H. Zhang. Smart collaborative automation for receive buffer control in multipath industrial networks. *IEEE Transactions on Industrial Informatics*, 16(2):1385–1394, 2020.

[16] F. Song, L. Li, I. You, and H. Zhang. Enabling heterogeneous deterministic networks with smart collaborative theory. *IEEE Network*, 35(3):64–71, 2021.

[17] F. Song, Y.-T. Zhou, Y. Wang, T.-M. Zhao, I. You, and H.-K. Zhang. Smart collaborative distribution for privacy enhancement in moving target defense. *Information Sciences*, 479:593–606, 2019.

[18] F. Song, M. Zhu, Y. Zhou, I. You, and H. Zhang. Smart collaborative tracking for ubiquitous power iot in edge-cloud interplay domain. *IEEE Internet of Things Journal*, 7(7):6046–6055, 2020.

[19] Vapnik and N. V. An overview of statistical learning theory. *IEEE Transactions on Neural Networks and Learning Systems*, 10(5):988–999, 1999.

[20] Z. Wenqing, Y. Hai, and S. Xuqiang. Target detection based on improved non maximum suppression algorithm. *Journal of Image and Graphics*, 23(11):10, 2018.

[21] C. Xiaolu. *Research on pedestrian detection based on classifier algorithm*. PhD thesis, University of Electronic Science and Technology of China, 2015.

[22] H. Xiting. *Introduction to Psychology*. Introduction to Psychology, 1991.

[23] L. Yacheng. *Research on Gaussian blurred image restoration and matching algorith*. PhD thesis, Huazhong University of Science and Technology.

[24] S. Yi. *Research and implementation of personnel type identification method in construction site*. PhD thesis, University Of Electronic Science And Technology Of China, 2020.

[25] L. Yugong. *Improvement of K-means clustering algorithm and its application in color image segmentation*. PhD thesis, Anhui University, 2018.

Author Biography

Chen Guangshuo is a postgraduate student at the School of Electronic and Information Engineering, Beijing Jiaotong University, China. His current research interests include communication technology and software engineering.