**Tech Science Press**

# Pedestrian Crossing Detection Based on HOG and SVM

## Yunzuo Zhang[*], Kaina Guo, Wei Guo, Jiayu Zhang and Yi Li

School of Information Science and Technology, Shijiazhuang Tiedao University, Shijiazhuang, 050043, China
[*]Corresponding Author: Yunzuo Zhang. Email: zhangyunzuo888@sina.com

**Abstract:** In recent years, pedestrian detection is a hot research topic in the field of computer vision and artificial intelligence, it is widely used in the field of security and pedestrian analysis. However, due to a large amount of calculation in the traditional pedestrian detection technology, the speed of many systems for pedestrian recognition is very limited. But in some restricted areas, such as construction hazardous areas, real-time detection of pedestrians and cross-border behaviors is required. To more conveniently and efficiently detect whether there are pedestrians in the restricted area and cross-border behavior, this paper proposes a pedestrian cross-border detection method based on HOG (Histogram of Oriented Gradient) and SVM (Support Vector Machine). This method extracts the moving target through the GMM (Gaussian Mixture Model) background modeling and then extracts the characteristics of the moving target through gradient HOG. Finally, it uses SVM training to distinguish pedestrians from non-pedestrians, completes the detection of pedestrians, and labels the targets. The test results show that only the HOG feature extraction of the candidate area can greatly reduce the amount of calculation and reduce the time of feature extraction, eliminate background interference, thereby improving the efficiency of detection, and can be applied to occasions with real-time requirements.

**Keywords:** Pedestrian detection; HOG; SVM; GMM

## 1 Introduction

With the continuous advancement of computer intelligence technology, it has become possible for the monitoring system to develop towards intelligence. At the same time, the large-scale application of computer image processing and information technology has gradually combined with video surveillance systems. How to use computer vision technology to process surveillance video effectively and in real-time becomes particularly important. If the video surveillance system can automatically detect abnormalities and alert people when abnormal behaviors occur, it will greatly improve people's work efficiency and find abnormalities early and deal with them. This research mainly used background extraction and pedestrian detection to analyze whether pedestrians are entering the dangerous warning area and cross-border behavior in the video.

Since 2005, Dalal et al. [1] proposed HOG, pedestrian detection technology has entered a stage of rapid development. The HOG feature is used to describe the gradient information of the pixel in the image. This feature fully describes the edge information of pedestrians and is not sensitive to changes in illumination. It is still the most widely used feature operator in the field of pedestrian detection.

The rate of pedestrian detection is easily affected by factors such as the dimension of the feature descriptor and the size of the detected images. The higher the dimensionality and the larger the image, the detection time will be longer [2]. The GMM has a good detection effect on overall human behavior. This method is mainly aimed at representing complex background scenes that the single Gaussian model that

cannot represent effectively, such as leaf shaking, water ripples, etc. [3–6]. GMM is often used in background modeling in complex scenes. Because it can adapt to changes in the background, it is better than other traditional models. So this article uses GMM to model the background. Therefore, firstly, extract the moving target, detect the moving area in the video [7–8], and then perform pedestrian detection, which can greatly reduce the feature detection area, thereby reducing the detection time and improving the detection efficiency. Therefore, this article proposes a method to detect moving targets through background modeling use GMM, which can reduce the interference of the background, and then use the method of HOG and SVM to classify the moving targets in the video, and finally detect the presence in the forbidden area of pedestrians are marked. The flowchart of this method is shown in Fig. 1 below.
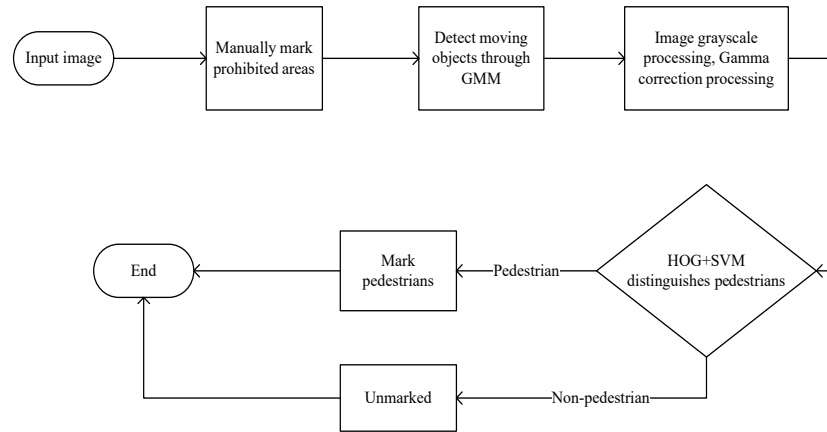


**Figure 1:** Pedestrian detection flowchart

## 2 Moving Target Detection

The main purpose of background modeling is to convert the moving target detection problem of sequence images into a two-classification problem based on the current background estimation, which divides all pixels of the input sequence into background and motion foreground, and then processes the classification results [9–10].

First, process the generated color image and observe the random variable $x$. The pixel sample of the variable at time $t$ is $x_t = \{r_t, g_t, b_t\}$, and the single sampling point $x_t$ taken by it satisfies the mixed Gaussian distribution probability density function, as shown in formula (1).

$$
\begin{cases}
p(x_t) = \sum_{i=1}^{k} w_{i,t} \times \eta(x_t, \mu_{i,t}, \tau_{i,t}) \\
\eta(x_t, \mu_{i,t}, \tau_{i,t}) = \dfrac{1}{\left|\tau_{i,t}\right|^{\frac{1}{2}}} e^{\frac{1}{2}(x_t - \mu_{i,t})^\tau \tau_{i,t}^{-1}(x_t - \mu_{i,t})} \\
\tau_{i,t} = \delta_{i,t}^2 I
\end{cases}
\tag{1}
$$

For the convenience of calculation, it is generally assumed that the channels of the pixels are independent of each other and have the same variance, to establish a Gaussian mixture model for each pixel. The effect diagram of moving target detection using Gaussian mixture background modeling is as follows. Fig. 2(a) is an image of the walking behavior video, Fig. 2(b) is the result of moving targets detected by using Gaussian mixture background modeling.

(a) Original image                                (b) The result of GMM

**Figure 2:** Target detection effect diagram

Gaussian mixture background modeling is a background representation method based on the statistical information of pixel samples. It uses statistical information such as the probability density of a large number of sample values of pixels in a long time to represent the background and then uses the statistical difference to determine the target pixel, which can perform complex dynamic background Modeling [11].

## 3 Pedestrian Detection Based on HOG Feature Combined with SVM Classifier

This article uses HOG and SVM methods to detect pedestrians in a video. The flow chart is shown in Fig. 3.
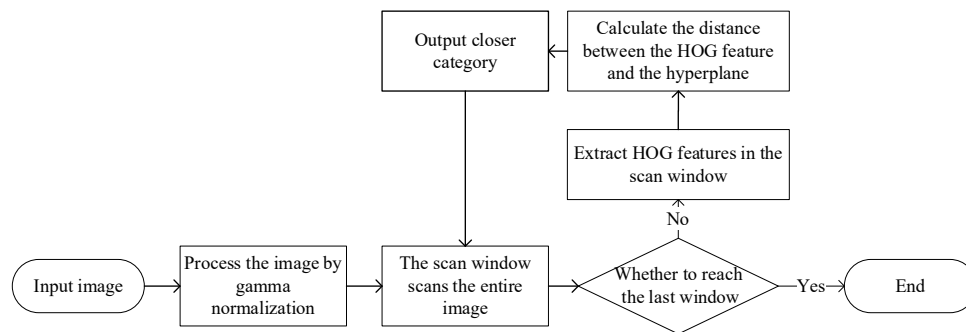


**Figure 3:** Hog combined with SVM pedestrian detection flow chart

### 3.1 HOG Features

HOG feature descriptors are commonly used in computer vision and image processing for object detection [12]. This feature is a histogram of the gradient direction of the local area of the image obtained by statistics, which represents the local gradient direction and the gradient intensity distribution characteristics. In this method, when the specific position of the edge is unknown, the distribution of the edge direction can also represent the contour of the pedestrian target.

First, the image needs to be preprocessed, the image to be detected is grayed out, and the Gamma correction method is used to normalize the color space of the input image to adjust the contrast of the image, reduce the impact of local shadows and light changes in the image, and suppress the noise interference of the image [13].

Then get edge gradient information. The gradient information mainly exists on the edge of the image, and the statistical information of the gradient can better describe the appearance and shape of the local target. The Sobel operator detects the edge based on the gray-scale weighted difference of the upper and lower, left and right adjacent points of the pixel, and reaches the extreme value at the edge. Sobel operator has a smoothing effect on noise and provides more accurate edge direction information, but the edge positioning accuracy is not high enough. The Sobel convolution factor is shown in formula (2), among them, the calculation of the horizontal gradient detection vertical edge is the Sobel convolution factor $G_x$, and the calculation of the vertical gradient detection horizontal edge dimension is the Sobel convolution factor $G_y$.

$$G_x = \begin{bmatrix} -1 & 0 & -1 \\ -2 & 0 & -2 \\ -1 & 0 & -1 \end{bmatrix}, G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{2}$$

Dalal et al. [1] proposed to use a $3 \times 3$ Sobel operator to perform gradient operations, and then calculate the gradient magnitude and direction by formula (3).

$$\begin{cases} G = \sqrt{G_x^2 + G_y^2} \\ \theta = \arctan \dfrac{G_y}{G_x} \end{cases} \tag{3}$$

The calculation of HOG is based on the density matrix of the uniform space to improve accuracy. When calculating the HOG features, as shown in Fig. 4, the feature descriptors of all cells in a block are connected in series to obtain the descriptor of the block, and the HOG of all blocks in the image The feature descriptors can be connected in series to get the HOG feature descriptor of the image, which is the final feature vector for classification. The calculation of HOG is based on the density matrix of the consistent space to improve accuracy. When calculating HOG features, as shown in Fig. 4 below, the feature descriptors of all cells in a block are connected in series to obtain the feature descriptors of the block. The block feature descriptors can be connected in series to obtain the HOG feature descriptor of the image, which is the final feature vector for classification.
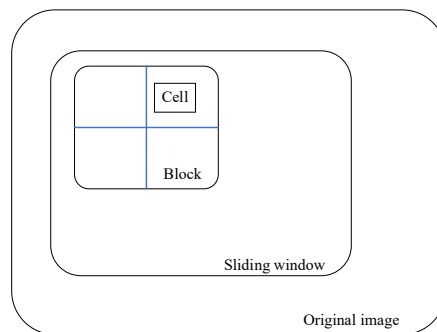


**Figure 4:** HOG feature extraction schematic

The relevant parameters of this experiment are set as follows: the cell unit is pixels, the block unit is the cell, the sliding window uses a fixed size, and the scan step is 8 pixels.

Finally, the contrast is normalized. Calculate the density of each histogram in this block, and normalize each cell unit in the interval according to this density. After normalization, better results can be obtained for lighting changes and shadows.

Compared with other characterization methods, HOG has many advantages. First of all, because the

calculation is performed on local grid cells, HOG can maintain good invariance to image geometric and optical deformation. Secondly, because of operations such as coarse spatial sampling, fine direction sampling, and local optical normalization, this method allows pedestrians to maintain an upright posture and have some subtle movements of their limbs. These subtle movements can be ignored and it will not affect the detection effect. Therefore, the HOG feature is particularly suitable for detecting people in images.

### 3.2 SVM Classifier

SVM is a two-category model. The learning strategy of SVM is to maximize the interval, which can be formalized as a problem of solving convex quadratic programming, which is also equivalent to the problem of minimizing the regularized hinge loss function. The learning algorithm of SVM is the optimal algorithm for solving convex quadratic programming.

The basic idea of SVM learning is to solve the separation hyperplane that can correctly divide the training data set and have the largest geometric interval. For a linearly separable data set, there are infinitely many such hyperplanes, but the separating hyperplane with the largest geometric interval is unique. Therefore, our goal is to find a segmentation plane that satisfies the classification requirements when the given sample point is separable, and to make the given sample point as far as possible from the segmentation plane [14]. That is, while ensuring the classification accuracy, maximize the blank area on both sides of the segmentation plane, because the larger the blank area, the better the generalization performance.

Use Python to complete the training of the SVM classifier. The first training result is not ideal, and the accuracy of image detection is average. When the human body occlusion is serious and the background is more complex, some misdetections and missed detections are shown in Fig. 5(a) below, because the number of negative samples is not enough and the diversity is not good enough. Next, using the detector that comes with OpenCV, you can clearly feel that the accuracy has improved a lot during the second training. Next, the bootstrap method is used to extract the difficult samples from the false positives, and add them to the negative samples [15–16]. Using the INRIA data set, the final SVM training effect achieved is shown in Fig. 5(b).
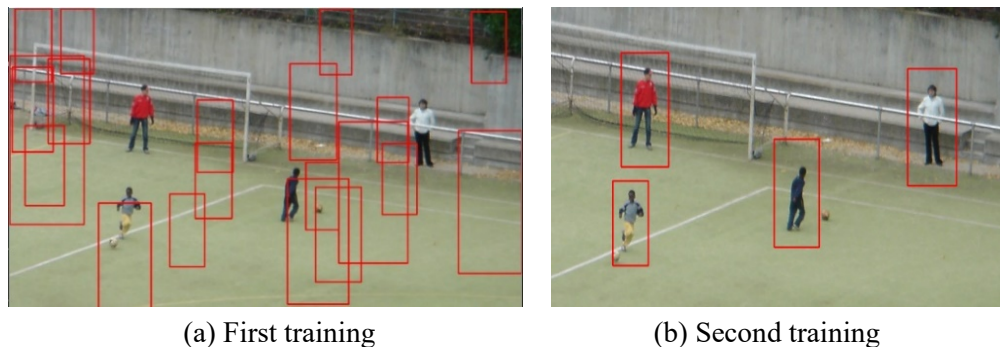


(a) First training                                         (b) Second training

**Figure 5:** Pedestrian detection effect

### 3.3 Combination of SVM and HOG Models

This article uses positive and negative samples from the French INRIA pedestrian dataset to train the SVM classifier. This data set is the upright pedestrians in the images and videos collected by Dalal during HOG pedestrian detection [17–18]. The data set contains positive samples and negative samples of pedestrian images. The images contained in the data set are relatively clear and friendly to training.

Detection process: Scan the input frame image according to a window of a given size, and extract the HOG feature value in the scan window. According to the SVM classification algorithm, the distance between the HOG feature of the scan window and the support-vector of the positive and negative samples determines the type of the scan window, and positive samples are detected [19].

When implementing pedestrian detection, this paper uses the HOG set SVM detector and detect

functions in Python to recognize pedestrians, and uses non-maximum suppression methods to optimize overlapping detection areas.

## 4 Experiment Analysis

This experiment uses the INRIA data set for experimental analysis. The INRIA data set has diverse pedestrian poses and complex and changeable backgrounds. In the field of pedestrian detection, this data set is highly recognized [20]. The number of positive samples in the training samples is 2416, and the number of negative samples is 453. The operating environment of the hardware is the Windows10 operating system, Intel(R) Core(TM) i5-8265U CPU, 16 G memory. The software platform is JetBrains PyCharm Community Edition 2019.

When performing pedestrian detection on images in the INRIA data set, the SVM+HOG in this article has already demonstrated superior detection results. Some experimental screenshot is shown in Fig. 6 below.

**Figure 6:** Some pedestrian detection result

Although there are some false detections and missed detections, it also shows excellent results. For the recognition rate of Haar feature and LBP feature [12], which are also detected in the INRIA data set. The experimental results are compared with the above common features combined with the SVM classifier. It has an advanced performance. The specific values are shown in Tab. 1.

**Table 1:** Comparison of different features combined with SVM classifier

| Feature type | Recognition rate(%) | Detection time(ms) |
|---|---|---|
| Haar | 77.4 | 53.27 |
| LBP | 83.64 | 49.01 |
| HOG | 90.78 | 25.61 |

In the process of pedestrian detection on the video, some fames may have the results shown in Fig. 7. In order to reduce the false detection of the background and reduce the amount of calculation, the method in this paper uses GMM to model the video. The modeling effect in different scenes is shown in Fig. 8, and then the modeled video is processed to obtain the video shown in Fig. 9. The result of this processing makes the video eliminate the interference of the background, and will not cause false detection of the background. Then use HOG+SVM for pedestrian detection on the modeled video, after the final processing, a screenshot of the video used for detection is shown in Fig. 9, and the recognition result is shown in Fig. 10.
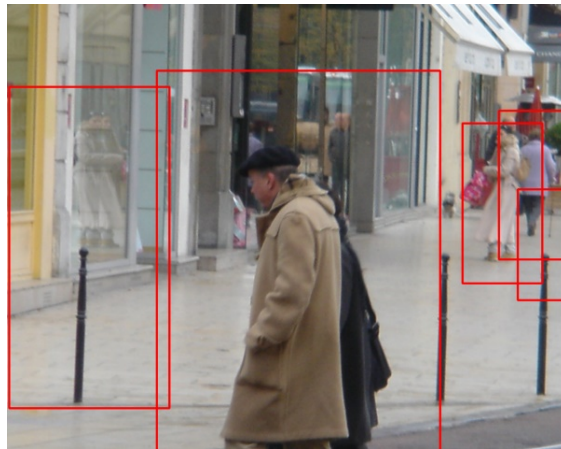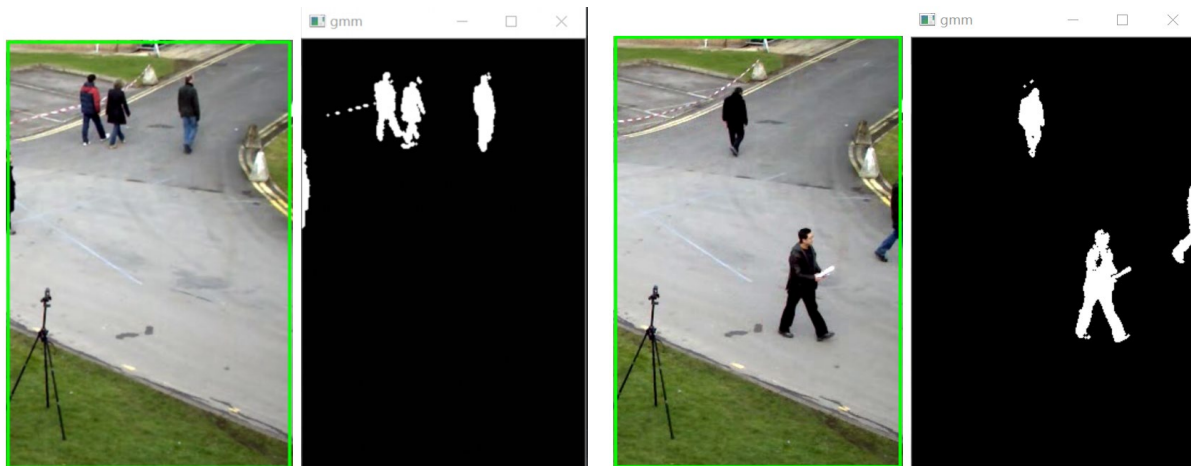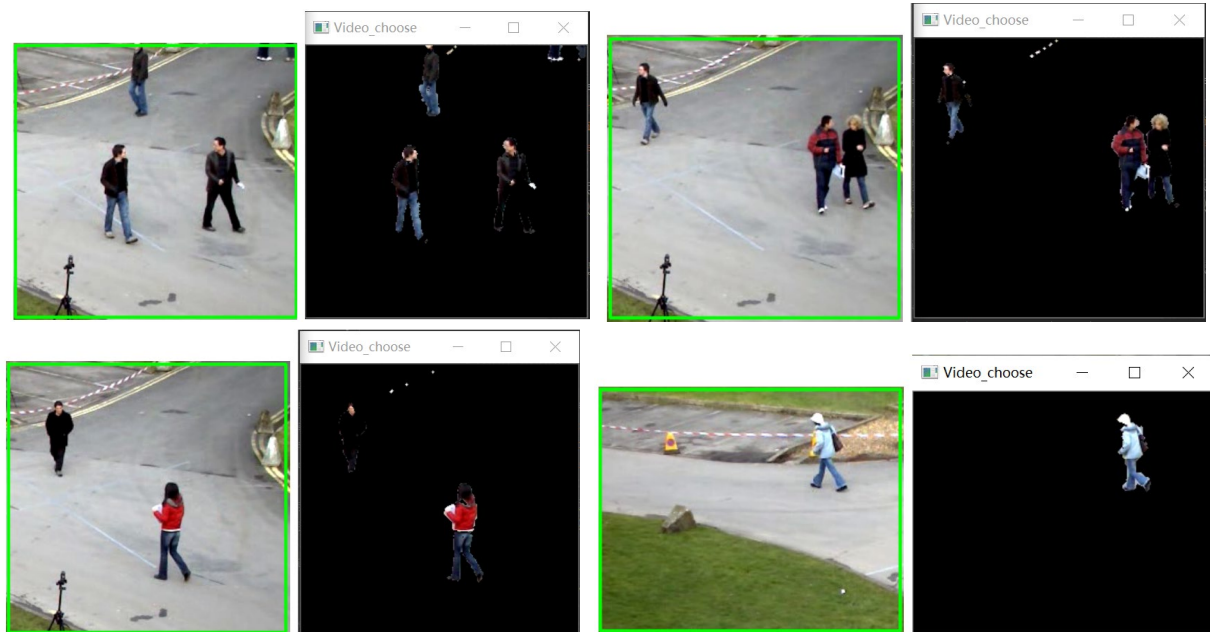


**Figure 7:** Pedestrian detection under background interference



(a) Scene 1

(b) Scene 2

**Figure 8:** The modeling effect in different scenes



**Figure 9:** Video for pedestrian detection

Experiment shows that the algorithm proposed in this paper can improve the detection efficiency, and the use of GMM modeling method can eliminate background interference and no longer cause false detection of the background. It can basically fulfill the requirements of fluency.
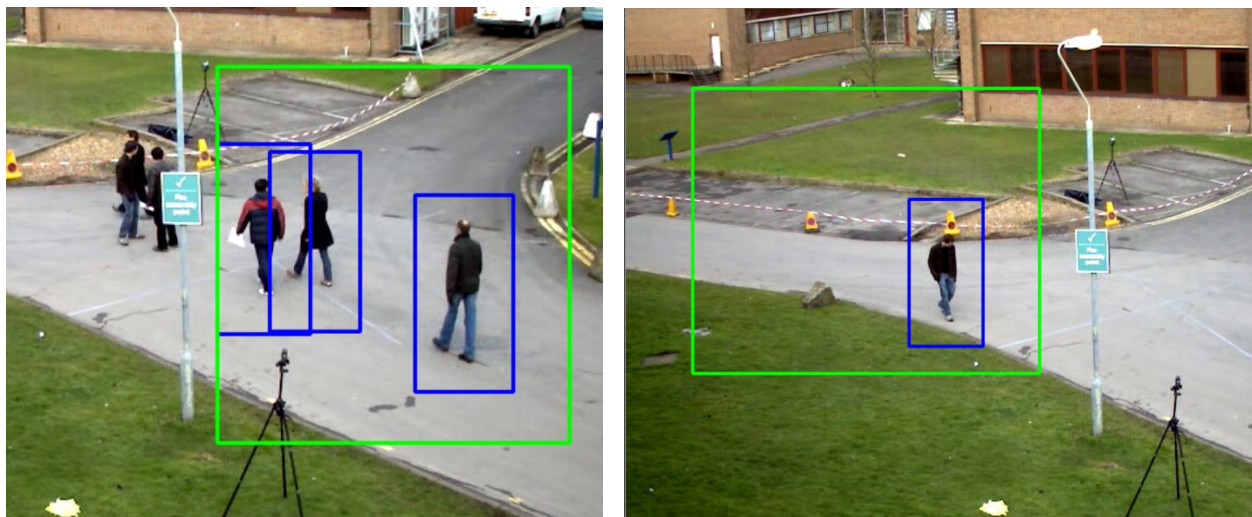
**Figure 10:** Pedestrian detection results

## 5 Conclusion

Aiming at the problem of pedestrian detection in intelligent surveillance systems, this paper proposes a method to detect whether pedestrians in the video are out of bounds. This method first performs a mixed Gaussian background modeling on the video frame, and then uses HOG+SVM to classify and detect pedestrians. In this way, background interference is eliminated during detection, the false detection rate is reduced, and the detection speed and efficiency are improved.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

[1] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. CVPR'05*, San Diego, CA, USA, pp. 886–893, 2005.

[2] X. P. Guo, J. S. Du, J. J. Bai and W. Wang, "Fast human detection algorithm based on BING objectness", *Application Research of Computers*, vol. 35, no. 11, pp. 3458–3461, 2018.

[3] Z. B. Zhao, "Video background modeling based on improved Guassian mixture model algorithm," M.S. thesis, Nanjing University of Posts and Telecommunications, Nanjing, 2013.

[4] X. L. Luo, "Background modeling in video tracking," M.S. thesis, National University of Defense Technology, Changsha, 2007.

[5] W. Liu, X. L. Hao and J. L. Lv, "Efficient moving targets detection based on adaptive Gaussian mixture modelling," *Journal of Image and Graphics*, vol. 25, no. 1, pp. 113–125, 2020.

[6] M. M. Lv and J. M. Sun, "Moving image target detection based on modified gaussian mixture model," *Semiconductor Optoelectronics*, vol. 40, no. 6, pp. 874–878, 2019.

[7] X. R. Mi and J. F. Cao, "Robust target tracking based on locality sensitive histogram," *Computer Engineering and Design*, vol. 39, no. 11, pp. 3526–3530, 2018.

[8] Y. C. Dong, Y. G. Shan and J. Yuan, "Pedestrian detection based on improved HOG," *Computer Engineering and Design*, vol. 41, no. 10, pp. 2921–2926, 2020.

[9] F. F. Qu, Z. F. Wang and J. J. Tian, "Video detection algorithm for object in dangerous area of mine based on background registration," *Industry and Mine Automation*, vol. 37, no. 3, pp. 48–50, 2011.

[10] S. W. Xia and X. Zhou, "Object tracking algorithm based on optical flow method and block training," *Computer Engineering and Design*, vol. 41, no. 7, pp. 2063–2068, 2020.

[11] W. J. Zhu, G. L. Wang, J. Tian, Z. T. Qiao *et al.,* "Spatio-temporal adaptive mixture of gaussians for moving

objects detection in complex background scenes," *Transactions of Beijing Institute of Technology*, vol. 38, no. 2, pp. 165–172, 2018.

[12] W. S. Li, Y. Han, M. H. Ruan and Z. X. Wang, "Improved pedestrian detection method based on enhanced HOG," *Computer Systems & Applications*, vol. 29, no. 10, pp. 199–204, 2020.

[13] H. Zhang, J. Wang, J. J. Shan and J. J. Shen, "Research on pedestrian detection algorithms based on HOG+SVM," *Journal of Baicheng Normal University*, vol. 34, no. 2, pp. 30–34, 2020.

[14] S. Naila, J. J. Yu, N. Yang, H. Kashif, J. Tang *et al.,* "A rapid recognition method for rice false smut based on HOG features and SVM classification," *Journal of Physics: Conference Series*, vol. 1576, no. 1, 2020.

[15] L. M. Gong, M. H. Xu, D. J. Xu and F. Y. Zhang, "Novel model of pedestrian detection based on Gaussian mixture model and HOG+SVM," *Journal of Shanghai University (Natural Science Edition)*, vol. 24, no. 3, pp. 341–351, 2018.

[16] F. Bi, W. Chen, W. Fang and X. Miao, "Fire detection method based on improved fruit fly optimization-based SVM," *Computers, Materials & Continua*, vol. 62, no. 1, pp. 199–216, 2020.

[17] R. Li, Y. Liu, Y. Qiao, T. Ma, B. Wang *et al.,* "Street-level landmarks acquisition based on SVM classifiers," *Computers, Materials & Continua*, vol. 59, no. 2, pp. 591–606, 2019.

[18] Z. Wang, R. Jiao and H. Jiang, "Emotion recognition using WT-SVM in human-computer interaction," *Journal of New Media*, vol. 2, no. 3, pp. 121–130, 2020.

[19] G. Yang, M. Yang, S. Salam and J. Zeng, "Research on protecting information security based on the method of hierarchical classification in the era of big data," *Journal of Cyber Security*, vol. 1, no. 1, pp. 19–28, 2019.

[20] Z. Liu, W. Gu and J. Xia, "Review of access control model," *Journal of Cyber Security*, vol. 1, no. 1, pp. 43–50, 2019.