Electronics, Communications and Networks
A.J. Tallón-Ballesteros et al. (Eds.)

doi:10.3233/FAIA231210

A New Pedestrian Detection Method Based on Histogram of Oriented Gradients and Support Vector Data Description

Lei KE and Yujing LUO[1]
School of intelligent manufacturing, Panzhihua University, Panzhihua City Sichuan Province 617000, China

Abstract. Auto-driving is an important technology development direction, and pedestrian detection is an indispensable key technology to achieve this goal, which belongs to the hot research field of computer vision. The main methods in the field of pedestrian detection are based on statistical learning, but in practice, this method has some problems such as incomplete negative sample data set and complex classifier. To solve these problems, a support vector domain data description and Histogram of oriented gradients pedestrian detection algorithm are proposed, hog and SVDD were used to learn and extract the features of the data set samples, and the penalty terms were solved by the equation. Joseph-louis Lagrange operator was introduced to reduce the complexity of the classifier, furthermore, the generalization degree between the training set of negative samples and the real scene is enhanced. First, a multi-scale candidate sample is generated by using the sliding window. Secondly, Histogram of oriented gradients statistics is used to extract the features of the image samples. Thirdly, based on the complete positive samples and the support vector domain, a single classifier with positive samples is trained to detect pedestrians. Experiments show that compared with the traditional pedestrian detection method based on statistical learning, the Hog and SVDD pedestrian detection algorithm has lower false positive rate and less total training samples, it has good generalization ability and certain practical research value, but needs more positive samples for training.

Keywords. Hog feature extraction, support vector network, pedestrian detection, summary learning

1. Introduction

Pedestrian identification is widely used in intelligent city and digital traffic, and it is also a significant area of exploration in image recognition. At present, with Self-driving technology, the automatic application of pedestrian detector algorithm has been widely accepted.

As a unique branch of object detection research, pedestrian detection has high difficulty. The difficulty degree of object detection mainly depends on the difference between object classes. The uniqueness of pedestrian detection has three aspects: first, it has a wide range of potential value in practice; second, the human body can adapt to a variety of environments and postures. This characteristic makes the human body

[1] Corresponding author: Yujing Luo, 706411754@qq.com.

excellent in motion and perception, so the research of pedestrian detection can be used for other object detection fields, the difference between pedestrians is large. Therefore, the study of hog feature extraction is actually to consider how to capture the feature to complete the monitor issue under the great environment change.

The Pedestrian identification algorithm is mostly accomplished by backdrop-based modeling method and summary learning method [1-3]. Background-based modeling is easily disturbed and affected by outside factors such as lighting and scenes, so Shandong is less robust. [4]. Yanhua Xu et al. [5] suggested the use of multi-scale Haar wavelet [6] and SVM [7] for Pedestrian identification, and suggested a sliding window method; Kim et al. [8] used cumulative scoring graph for examination, and used ADABOOST rule system for sort; Liu Ben et al [9] proposed HOG way to describe image specimens and use SVM for classification.

All the sorting machines in the above algorithms need to be trained with negative samples. However, in practical applications, the number of negative samples that can be used for training can not be applied to all scenarios, so the algorithm can not be generalized. To solve this problem, we often increase the number of negative samples to improve the robustness of the algorithm. However, there are drawbacks to this approach, which can make the model more complex and difficult to train and solve.

This paper proposes a pedestrian identification algorithm based on HOG [9] and sustain vector domain data [10]. The proposed algorithm generates regional suggested samples through multiple scales of sliding windows, samples are sampled and histogram statistics in the gradient direction are made. Based on these features, a full set of positive samples is trained to generate a single classifier for pedestrian recognition. Importantly, the algorithm does not want to lead a negative sample set. It reduces the influence of the incomplete negative specimens on the sorting machine's ability.

Ultimatly, a great deal of contrast tests is carried out. Compared with the classical algorithm based on HOG and SVM, the proposed algorithm needs more positive specimens in drilling when its performance is comparable, however, the overall sample size required is much smaller. Moreover, this algorithm is easier to avoid false detection.

2. Algorithm flowchart

The focal point of algorithm is to search for the similarity of positive specimens when the integrity of positive specimens is confirmed. The algorithm can be separated into two types: feature extraction and aggregate classes. From the direction of the implementation algorithm, it is separated into training section and application section. The positive sample set of manual marking is a necessary condition for the model. To train a single classifier based on support vector data description (SVDD), these samples need to be sampled to a uniform size (64×64) and their HOG features extracted.

By using a sliding window with variable size, we can extract candidate samples from the input image and preprocess these samples to extract HOG features. We then input these features into a pre-trained single classifier for classification. The overall flow chart of this algorithm is shown in Figure 1.

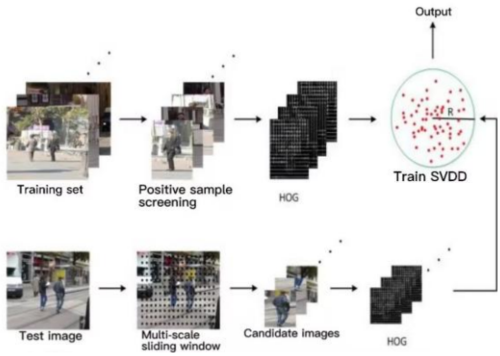

Figure 1. overall flow chart of the algorithm

3. Gradient direction histogram

There are four stages in the generation of gradient histogram, which are advance processing, gradient computation, Calculate the gradient and HOG descriptor production.

In the pre-processing stage, we use Gamma algorithm to adjust the illumination and contrast of the image to reduce their influence on the gradient direction histogram. This stage is one of the necessary steps to generate the HOG feature. The formula for Gamma's algorithm is below.

$$I(x,y)' = I(x,y)^{Gamma} \quad (1)$$

Where $I(x, y)$ is the pixel in the image and Gamma is typically 0.5.

Formula (2) is the formula for calculating the gradient.

$$\begin{aligned} G_x(x,y) &= I(x+1,y) - I(x-1,y) \\ G_y(x,y) &= I(x,y+1) - I(x,y-1) \\ G(x,y) &= \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \quad (2) \\ \alpha(x,y) &= \tan^{-1}(\frac{G_y(x,y)}{G_x(x,y)}) \end{aligned}$$

Where, $G(x,y)$ represents the size of the gradient and $\alpha(x,y)$ represents the direction of the gradient.

Then the local image (6 × 6) is encoded. The number of quantized directions is 9. Thus, a code representing the image features is generated in each local region. Integrating the directional gradient statistics of each local area, the whole image description is generated. Hog feature is an operation on the local part region of an image, which can maintain a certain degree of stability to the geometry and optical distortion of the image. Fig. 2 is a sketch of HOG feature extraction.

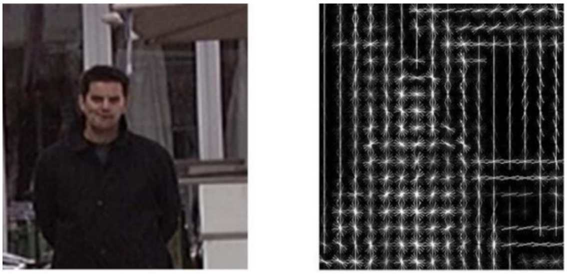

Figure 2. the sketch of HOG feature extraction

4. Support vector domain data description

The standby vector domain [10] can be seen as an aggregation algorithm, which trains a “Spherical” hyperplane on the basis of the training specimens, so that high-dimensional open air is divided into two sections, and therefore can be applied to sort. Different from the Sustain vector machine, the SVM data representation does not need negative specimens to participate in the drilling, but the positive sample boundary is trained. Therefore, the key of applying support vector domain data description in Practical application is to guarantee the integrity of positive samples.

4.1 *Support the training of vector domain data description*

Let $X=\{x_i,\ i=1,2,\ldots M\}$ Be a sample set whose $x_i \in R^N$ is both the perspective of the specimen N. If the center of the“Spherical” hyperplane is a and the radius is R, the hyperplane satisfies the formula

$$(x_i-a)(x_i-a)^T \le R^2 \qquad (3)$$

The training SVM can be considered to solve a limited improvement issue, and its mathematical expression is shown below(4)

$$\begin{aligned} \min \quad & F(R,\ a) = R^2 \\ s.t. \quad & R^2 \ge (x_i-a)(x_i-a)^T \end{aligned} \qquad (4)$$

Yet, in practice, noise is everywhere, so, on the foundation of the formula, we introduce a Punishment term ε_i , such as the formula(5)

$$\begin{aligned} \min \quad & F(R,\ a,\ \varepsilon) = R^2 + C\sum_{i=1}^{M}\varepsilon_i \\ s.t. \quad & R^2+\varepsilon_i \ge (x_i-a)(x_i-a)^T \\ & \varepsilon_i > 0 \end{aligned} \qquad (5)$$

Where C is the Punishment coefficient. Selecting KKT algorithm [11], Joseph-Louis Lagrange operator is lead, the formula can be transformed into formula solution.

$$L(R,\ a,\ \varepsilon,\ \alpha) = R^2 + C\sum_{i=1}^{M}\varepsilon_i$$
$$-\sum_{i=1}^{M}\alpha_i[R^2+\varepsilon_i-(x_i^2-2ax_i+a^2)]-\sum_{i=1}^{M}\gamma_i\varepsilon_i \quad (6)$$

Calculate the partial derivative of the upper formula and bring the result into the upper formula(7)

$$L(R,\ a,\ \varepsilon,\ \alpha)=\sum_{i=1}^{M}\alpha_i(x_i\bullet x_i)-\sum_{i=1}^{M}\sum_{j=1}^{M}\alpha_i\alpha_j(x_i\bullet x_j) \quad (7)$$

In realistic applications, we will introduce a kernel function to shine upon the nonlinearity of a low-dimensional linear space to Galway, and the formula will be transformed into the formula(8)

$$L(R,\ a,\ \varepsilon,\ \alpha)=\sum_{i=1}^{M}\alpha_iK(x_i\bullet x_i)$$
$$-\sum_{i=1}^{M}\sum_{j=1}^{M}\alpha_i\alpha_jK(x_i\bullet x_j) \quad (8)$$

The "Spherical" hyperplane can be obtained by solving the above formula and obtaining the optimal α_i to minimize the formula. The radius R can be acquired from α_i. As shown in Figure 3, the radius R is the final product.

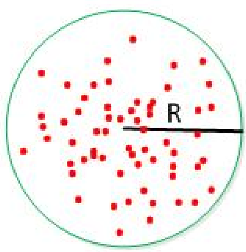

Figure 3. diagram of training result

In practical applications, we extract HOG features from positive sample sets and get $X=\{x_1,x_2,\ldots,x_M\}$ training sample set. In practice, there is no zero sample in X.According to the formula to obtain $A=\{\alpha_1,\alpha_2,\ldots,\alpha_M\}$, where A elements and X elements one-one corresponding. Suppose $X'=\{x_1,x_2,\ldots,x_N\}$, satisfying $\forall x_i\in X',\ 0<\alpha_i<C$, brings the elements of X' into the equation, and the radius R is:

$$R = 1 + \frac{1}{N}\sum_{n=1}^{N}(-2\sum_{i=1}^{M}(\alpha_i K(x_n, x_i))) + \sum_{i}^{M}\sum_{j}^{M}\alpha_i\alpha_j K(x_i, x_j) \tag{9}$$

4.2 *Support vector domain data descriptions are used for classification*

For the sample x_u to be classified, the distance between the sample x_u and the"Sphere center" of the hyperplane is determined to be positive if the formula is satisfied, otherwise, it is judged to be negative.

$$(x_u - a)(x_u - a)^T \leq R^2 \tag{10}$$

In practice, the formula is modified to obtain the following formula.

$$K(x_u, x_u) - 2\sum_{i=1}^{M}\alpha_i K(x_u, x_i) + \sum_{i=1}^{M}\sum_{j=1}^{M}\alpha_i\alpha_j K(x_i, x_j) - R^2 \leq 0 \tag{11}$$

5. Experiment

In order to verify the effectiveness of the algorithm, a series of experiments are carried out in the environment of MATLABR2013A. The computer performance parameters used in the experiment were: Intel (r) Core (tm) i5 -3470 CPU, 4GB ram, 64--bit operating setup. The image sources applied in the simulation are the MIT pedestrian detection data set and the data set INRIA.

5.1 *The results of this algorithm*

Figures 4 and 5 show two typical results. The red box is the correct pedestrian detection and the white box is the false detection. The positive sample of training set is 5000, and the positive sample of test set is 768. Table 1 shows the test results of the algorithm.

Table 1. outputted results of the algorithm test set

	Detecti on rate	False positive rate
HOG+SVDD	88.5%	1.1%

Figure 4. algorithm test results I

In general, the proposed algorithm can monitor the pedestrian correctly, and the location is relatively accurate. In addition, our algorithm can also fit to the size of the objective and the influence of distance.

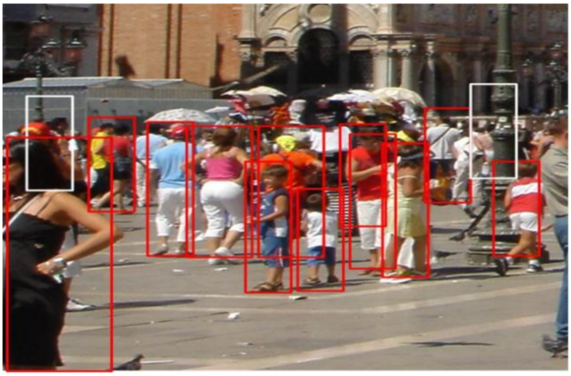

Figure 5. algorithm test results II

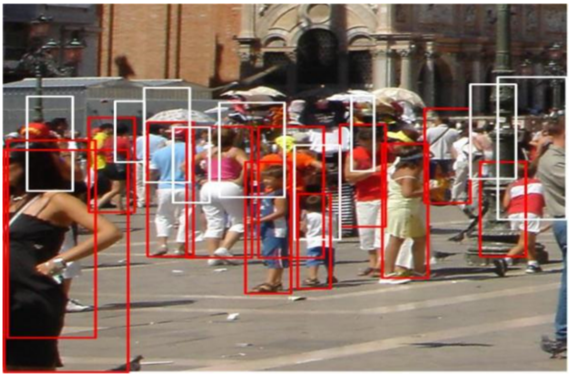

Figure 6. HOG + SVM algorithm test result II

Figure 5 shows a more crowded situation, while the survey outcomes of the classical algorithm under the same picture are shown in Fig. 6. The algorithm can detect most pedestrians with fewer false positives. However, in the dense pedestrian areas, especially in the overlap, there are multiple detection of the same pedestrian and the detection of the overlap of pedestrians as the same pedestrian situation. In addition, it needs to be explained that the algorithm in the face of partially occluded pedestrians, there is also a phenomenon of missed detection. Under the condition that the training set is relatively complete, there is misdetection in the algorithm, but the misdetection rate is relatively ideal.

5.2 *Contrast test*

Using the same data set, we compare our algorithm with the classical HOG + SVM method. The purpose of this experiment is to compare the training samples required by the two algorithms with the same performance.

Table 2. experimental comparison results

	Detection rate	False positive rate	Training positive specimens	Training specimens
SVM + HOG	70 .2%	7 .8%	1000	6000
SVDD + HOG	72 .5%	5 .8%	2000	2000
SVM + HOG	83 .2%	4 .8%	2000	22000
SVDD + HOG	82 .7%	3 .1%	3500	3500
SVM + HOG	89 .2%	2 .8%	2400	23000
SVDD + HOG	88 .5%	1 .1%	5000	5000

Table 2 shows the results of comparative experiments. The table shows that the algorithm performs well in controlling false testing, and the false detection rate is lower when the testing ratio is equivalent to the classical algorithm. Under the condition of equivalent property, the algorithm needs more positive samples, but less training samples.

In order to further study the specific performance index of the proposed algorithm, the method proposed in literature [12-14] is used to compare the experimental results.

Table 3. Comparison of experimental results of different algorithms

	Detection rate	False positive rate	Training samples
Ref. 12	86.3%	2.6%	16000
Ref. 13	87.3%	5.1%	12000
Ref. 14	88.9%	7.1%	8000
HOG+SVM	89.2%	2.8%	23000
This article methods	88.5%	1.1%	5000

Table 3 shows that the detection rate of this method is in the middle level, but the performance of the proposed method is better than that of other methods in terms of false detection rate and training sample size, it effectively alleviates the problem of too many training data samples in the territory of machine vision and the high error detection rate caused by over-fitting training.

6. Conclusion

Pedestrian detection is a key first step in intelligent recognition, which is used to determine whether there is a pedestrian in the input image and provide corresponding position information. Because the appearance of pedestrian is easily affected by dress, size, occlusion, posture and visual angle, pedestrian detection has become a hot and difficult research area of computer.

In order to settle the perplex of poor generalization performance due to incomplete training of negative specimens in pedestrian monitor algorithm using statistical studying means, in this paper, we propose a support vector domain data description and HOG-based pedestrian detection algorithm. First, multi-scale samples of candidate areas are generated, then histogram statistics of gradient directions are performed, and finally, a single classifier is trained to complete pedestrian detection using a complete set of positive samples. Compared with the classical svm+hog algorithm, the Hog + SVM algorithm has higher wrong monitor rate and needs more samples, so it is more advantageous to detect the pedestrian quickly and accurately. The complexity of the classifier is reduced, and the generalization between the training set of negative samples and the real scene is further enhanced. Due to the adoption of SVDD feature classification and training method, the number of SVM kernel functions is declined, the over-fitting phenomenon in the classification process and training process is reduced, and the false detection rate of the algorithm is declined to a certain circumstance, however, there is also a partial lack of fitting, which is one of the main problems of machine learning algorithms.

Hog feature dimension is too great and there are a lot of verbose traits in the algorithm, which leads to SVM training difficulty and slow algorithm. In the future, we can find more efficient and faster trait collection ways to solve the above problems and better deal with the pedestrian detection problem. Solving these problems is of great significance for realizing efficient pedestrian detection algorithm and popularizing its application. Among them, the computation and extraction of HOG feature data needs further research on the gradient computation and region coding algorithm, to find a more efficient method to replace the current method, in order to reduce the redundancy in the process of feature extraction, for example, subspace clustering algorithm combined with HOG algorithm for feature information extraction, but how to achieve the combination of the advantages of both still has some problems, this is one of the following research directions. In the SVDD algorithm, it is also necessary to distinguish some data whose feature information is close to each other in order to reduce the under-fitting of the training results.

References

[1] Monty Stephanie, Rigaut François, McDermid Richard, Baumgardt Holger, Cranney Jesse, Agapito Guido, Mendel J Trevor, Plantet Cédric, Greggio Davide, Stetson Peter B, Fiorentino Giuliana, Haynes Dionne. Towards realistic modelling of the astrometric capabilities of MCAO systems: detecting an intermediate-mass black hole with MAVIS[J]. Monthly Notices of the Royal Astronomical Society,2021,507(2).

[2] Mengru Ren, Honglu Hou, Xiulai Han. Gabor Pedestrian detection based on fast HOG feature[J].Computer system applications,2021,30(10):259-263.DOI:10.15888/j.cnki.csa.008136.

[3] Pingping Tong, Linlin Gao, Jiaman Gu, Xinyu Zhang, Xiaoxian Zhu. Research on the early warning method of automobile assistant driving system[J].Journal of Changshu Institute of Technology,2023,37(02):120-124.DOI:10.16101/j.cnki.cn32-1749/z.2023.02.017.

[4] Danyi Jin. Research and development of long-distance pedestrian detection system software based on background modeling[D].Zhejiang University,2021.DOI:10.27461/d.cnki.gzjdx.2021.001532.

[5] Yanhua Xu, Rongya Zhou. Application of Haar wavelet numerical solution principle in image edge detection algorithm[J].Electronic design engineering,2022,30(14):110-114.DOI:10.14022/j.issn1674-6236.2022.14.024.

[6] Bosl William J.,Loddenkemper Tobias,Vieluf Solveig. Coarse-graining and the Haar wavelet transform for multiscale analysis[J]. Bioelectronic Medicine,2022,8(1).

[7] Ramakrishnan Rahul O.,Friedrich Benjamin M.. Learning run-and-tumble chemotaxis with support vector machines[J]. Europhysics Letters,2023,142(4).

[8] Kim Donggeun, Park Sangwoo, Kang Donggoo, Paik Joonki. Real-time Robust Object Detection Using an Adjacent Feature Fusion-based Single Shot Multibox Detector[J]. IEIE Transactions on Smart Processing & Computing,2020,9(1).

[9] Liu Ben, Gao Feng, Zhao Pengxuan, Li Yan, Li Wenqiang. Marker Identification of Disc Workpiece Based on HOG-SVM Classifier[J]. Mobile Information Systems,2022,2022.

[10] Zheng Yunfei, Wang Shiyuan, Chen Badong. Robust one-class classification with support vector data description and mixed exponential loss function[J]. Engineering Applications of Artificial Intelligence,2023,122.

[11] Jidong Lai, Xian Yao, Jianhui Su, Cheng Chen. A method of DAB reflow power optimization based on KKT algorithm[J].Power Electronics Technology,2022,56(06):103-106.

[12] Pejun Sun, Zhongrong Zhang, Qiming Li, etc. Pedestrian detection method based on improved multi-scale residual network[J].Computer Engineering and design,2023,44(03):762-769.

[13] Xiaohong Wang, Zheqi Chen. Pedestrian detection in infrared image based on Yolov5 algorithm[J].Laser and infrared,2023,53(01):57-63.

[14] Dan Liu, Huilan Wang, Hanwen Zeng etc. Improved pedestrian detection algorithm for FCOS[J].Computer Engineering and design,2022,43(11):3264-3270.