

Research Article

Multistage Analysis of Abnormal Human Behavior in Complex Scenes

Lei Cai ¹, Peien Luo ², and Guangfu Zhou ²

¹School of Artificial Intelligence, Henan Institute of Science and Technology, Xinxiang 453003, China

²School of Information Engineering, Henan Institute of Science and Technology, Xinxiang 453003, China

Correspondence should be addressed to Lei Cai; cailei2014@126.com and Guangfu Zhou; zhouguangfuhn@163.com

Received 8 July 2019; Revised 25 September 2019; Accepted 30 September 2019; Published 15 November 2019

Academic Editor: Antonio Lazaro

Copyright © 2019 Lei Cai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Effective abnormal human behavior analysis serves as a warning signal before emergencies. However, most abnormal human behavior detections rely on manual monitoring at present. This method is criticized for being subjective and lack of timeliness. In response to the problems above, this paper proposes a multistage analysis method of abnormal human behavior in complex scenes. This paper firstly differentiates the abnormal behavior roughly from a large monitoring area with similarity measurement applied to the social force model, and precise analysis is conducted thereafter. The multistage analysis, based on the three-frame difference algorithm, is used for intrusion, left-behind baggage detection, and motion trajectory identification. The experimental result demonstrates the superiority of the proposed method in UMV, CAVIAR, and datasets. To demonstrate the adaptability and generalization ability of the proposed method, this paper selects the CVC and JAAD driving anomaly detection data sets to test the method. Experimental results show that the proposed method is superior to the existing methods.

1. Introduction

Behavior analysis, first introduced by psychology, utilizes psychological knowledge to describe, analyze, predict, and control an individual's external manifestations. The way human beings control his action can be divided into two levels: self-control and social control. The goal of self-control is adapting to the scene by reducing behavioral deficiencies or behavioral excesses. The so-called abnormal human behavior means that the behavior of an individual deviates from social norms, rules, or conventions. Analysis of abnormal behavior, which focuses on the causes, forms, and potential hazards, can contribute to the precaution of unusual events and loss mitigation.

Video security monitoring system uses video as a means of detection to monitor, control, and record information. When an unusual event occurs, the video surveillance system will call the image related to the alarm region to observe and reconfirm the condition on the spot. The current video surveillance system combines the latest research achievement

of computer vision, sensor networks, and so on. It could be used in various scenarios, including government, transportation hubs, hospitals, and schools. However, the current video surveillance system rarely has a prediction function and relies on personal judgment. Human emotions are complicated, so video surveillance is prone to subjectivity and lag of decision-making.

A single control model cannot meet current social needs. Multistage decision-making refers to the process in which problems can be decomposed into several interrelated stages chronically, and at each stage, decisions must be made. Therefore, the whole process is a decision sequence. The multistage model has strong reliability, robustness, and a wider range of applications.

To solve the above problems, this paper proposes a multistage analysis method of abnormal human behavior in complex scenes. Our approach works in three steps. Firstly, we make a rough analysis of the monitoring area by using the social force similarity measurement method. Through the interaction between social force and individual desire

force, the text differentiates the normal and abnormal behavior in the monitoring area. Secondly, we propose a multimodel human behavior analysis method based on the three-frame difference algorithm. At last, intrusion, left-behind baggage detection, and motion trajectory identification in the unusual region are detected specifically. The main contributions of this paper are as follows:

- (1) The similarity measure model is introduced into the social force model, reducing the amount of data calculation in behavior detection
- (2) A multimodel human behavior detection approach is established to improve the robustness of the analysis
- (3) A large number of experiments has been carried out to evaluate the performance of this approach. The result shows that our approach is superior to the existing methods

The rest of this paper is arranged as follows. The second part provides an overview of related work. The third part builds a social force similarity measurement model for human behavior region detection. The fourth part conducts a multimodel abnormal human behavior analysis. The fifth part analyzes the experimental results. Finally, the sixth part summarizes the paper.

2. Related Work

The correct judgment of abnormal human behavior analysis can serve as an effective warning signal before emergencies and hence minimize the damage caused by unexpected events. Many scholars have made outstanding contributions to the analysis of abnormal human behaviors in complex scenes. Lan et al. [1] proposed a multifeature sparse representation framework for visual tracking, in which multiple sparse models were decomposed to distinguish different features. Most CNN-based trackers use hierarchical features to represent targets. However, the layered feature is not completely effective in separating the target from the background, especially in complex interference environments such as heavy occlusion and illumination variation. To solve the above problems, Qi et al. [2] proposed a CNN-based tracking algorithm which hedges deep features from different CNN layers to better distinguish target objects and background clutters. Many experimental results show that the target-tracking method is effective. For autonomous driving application, traditional tracking methods usually distinguish between a target and a background by training a support vector machine (SVM). When the training sample is nonlinear, the tracker will not be able to track the target. To address this problem, Zhang et al. [3] proposed a distance-based distance learning method based on deep learning. Experiments have shown that even without model updates, the proposed method achieves favorable performance on challenging images. Zhang et al. [4] proposed a research scheme using density-based clustering and backtracking strategy to detect small dim targets. The algorithm extracted the trace

with the highest clutter suppression ratio, increasing the accuracy of target detection. Rotation invariance was introduced into aerial images of urban scenes by ELMikaty and Stathaki [5]. A feature detection method which analyzes the principle of wavelet transform applied for signal filtering was proposed in order to obtain high detection performance for small targets of complex multimedia images [6]. Asif et al. [7] presented an object segmentation approach for indoor scenes using perceptual grouping algorithm.

When it comes to establishing a multistage model, Grey relational analysis method and an objective optimization model based on maximum entropy are utilized to make the decision-making more comprehensive [8]. Tammi et al. [9] proposed that using the clustering techniques with the amalgamation of the neural network had higher accuracy in detecting the attacks. To address image forgery problem, an algorithm [10] was designed to classify the image blocks based on a feature present in multi-compressed JPEG images. Huang et al. [11] presented a bottom-up-based framework for salient object detection without any prior knowledge. The approach was more effective in highlighting the salient object and robust to background noise.

In terms of human behavior analysis, Iosifidis et al. [12] demonstrated a novel nonlinear subspace learning technique for class-specific data representation, which was obtained by applying nonlinear class-specific data projection to a discriminant feature space. Under the guidance of kinematics, Xiao et al. [13] proposed a data-driven approach to identify typical head motion patterns and stressed on analyzing human behavioral characteristics via signal processing methods. Human dynamics provides new ways of researching human behavior by exploiting the statistical analysis method. The interval time series and the number series of individuals' operating behavior was investigated by a modified multiscale entropy algorithm, which provided insights for further understanding of individual behavior at different time scales [14].

With respect to abnormal human behavior detection, pedestrian behavior modeling and analysis in video surveillance is important for crowd scene understanding. Yi et al. [15] proved the reliability in judging normal and abnormal pedestrian behavior by simulating pedestrian behavior and established two large pedestrian walking route datasets for future research. Hu et al. [16] proposed to detect abnormal driving, which may incur tragic consequence for individuals and the public, by analyzing normalized driving behavior. The driver's anomaly detection is very important in the field of automatic driving, which can prompt the driver to inform the danger. The detection of the driver is extremely difficult due to camera shake, a sharp change in the speed of the vehicle, and the like. In response to the above problems, Yuan et al. [17] proposed a spatial local constrained sparse coding method for anomaly detection of traffic scenes. From the experimental results, the proposed method is more efficient than the popular competitors and produces higher performance. By employing 3D discrete cosine transform of the target

in different frames, Yuan et al. [18] designed a multiobject tracker, which makes tracking analysis possible in high-density crowds. The experimental results of several public crowd video datasets verify the effectiveness of the proposed method. A taxonomy of novel classes of neighbor-based trajectory outlier definitions was designed to detect abnormal moving objects. The experimental result showed that the method worked well for high-volume trajectory streams in near real time [19].

To sum up, scholars have proposed many efficient solutions in the analysis of abnormal human behavior. However, there are still some problems waiting to be resolved: (1) manual monitoring is prone to subjectivity. (2) There is time lag in the discrimination of abnormal human behavior.

3. Human Behavior Region Detection—a Rough Analysis

In human behavior detection, human group behaviors are difficult to describe. There is no precise definition of abnormal human behavior. The so-called abnormal behavior is judged by the deviations from group behavior characteristics. In large-scale video surveillance, occlusion, low resolution, and interaction between human beings make it difficult to detect abnormal human behavior. LDA model [20, 21] can depend on prior category knowledge and experience in dimension reduction, distinguish normal and abnormal frames in the video by using threshold value. Unfortunately, the result will produce overfitting data. Therefore, this paper proposes an abnormal human behavior detection method based on social force similarity measurement model, which determines whether a person's behavior is abnormal according to the consistency with the surrounding scene. LDA-H adds similarity measure model Hash [22] to LDA to solve the problem of overfitting. Only two kinds of data—normal and abnormal behavior, are necessary for analyzing human behavior. LDA-H projects data into low dimensions, then uses a similarity measurement model to make the projection points of each data category as close as possible, and finally determined the normal and abnormal human behaviors.

First, each image $u(x)$ in the UMN dataset is pre-processed, with the image binarized and size adjusted to $100 * 100$. Secondly, feature extraction is carried out on $u(x)$ to obtain the feature vector $R = \{R_1, R_2, \dots, R_n\}$, where R_i represents a feature point vector in the graph. The compression of the feature matrix is achieved by summing the multiperspective feature matrices.

$$H(i) = \sum_{x=1}^n R_{i,x}. \quad (1)$$

This paper uses the social force model [23] to quantify H . In a certain scene, people usually do things with a specific purpose. Thus, the algorithm must take into account the action potential field model of every person presents in the scene. If a pedestrian's normal speed is

v_i and his expected speed is v_i^p , the model of personal desire force in the scene is defined as follows:

$$F_p = \frac{1}{\tau} (v_i^p - v_i), \quad (2)$$

where τ is the similarity threshold in the similarity metric model. The interaction force between individual behavior and the scene F_{int} could be divided into individual desire to force F_p and social influence F_w . Therefore, this paper defines the interaction force as follows:

$$F_{\text{int}} = F_p + F_w. \quad (3)$$

The influence of v_i^p on the similarity measurement of social forces will result in an opposite disturbance, i.e., the interaction force F_{int} will fluctuate and the abnormal behavior will happen. Thus, the individual abnormal behavior model is derived as

$$v_i^q = (1 - p_i)v_i^p + p_i \langle v_i^c \rangle, \quad (4)$$

where p_i is the constant parameter of the social force model in case of anomaly. $\langle v_i^c \rangle$ is a normal pedestrian's trend scalar. A similarity measurement model is associated with the social force model to measure the interaction force similarity of all human beings in the dataset.

The particle motion model of each pedestrian is defined as $v_i = O_{\text{ave}}(x_i, y_i)$, and the individual expected speed of particle motion is

$$v_i^q = (1 - p_i)O(x_i, y_i) + p_i O_{\text{ave}}(x_i, y_i). \quad (5)$$

Finally, based on a given scenario, the interaction force of each person in a group is estimated to be

$$F_{\text{int}} = \frac{1}{\tau} (v_i^q - v_i) - \frac{dv_i}{dt}. \quad (6)$$

In this paper, the image information is mapped to 0 or 1 according to the characteristic similarity threshold of human behavior τ : 1 and 0 stand for R_i 's normal behavior and abnormal behavior, respectively. Finally, the hash value H is obtained.

$$h(i) = \begin{cases} 1, & R_{i,x} \geq \tau, \\ 0, & R_{i,x} < \tau. \end{cases} \quad (7)$$

τ represents the motion threshold of similar particles of all objects. When human behavior characteristics are greater than or equal to τ , the model determines that the characteristics are similar and outputs the normal behavior. In contrast, when the human behavior characteristics are less than τ , the model determines that the characteristics are different and outputs abnormal behaviors, as shown in Figure 1.

3.1. Abnormal Human Behavior Detection—a Precise Analysis. In the detection of abnormal human behavior, a wide range of rough detection is way from enough. A precise

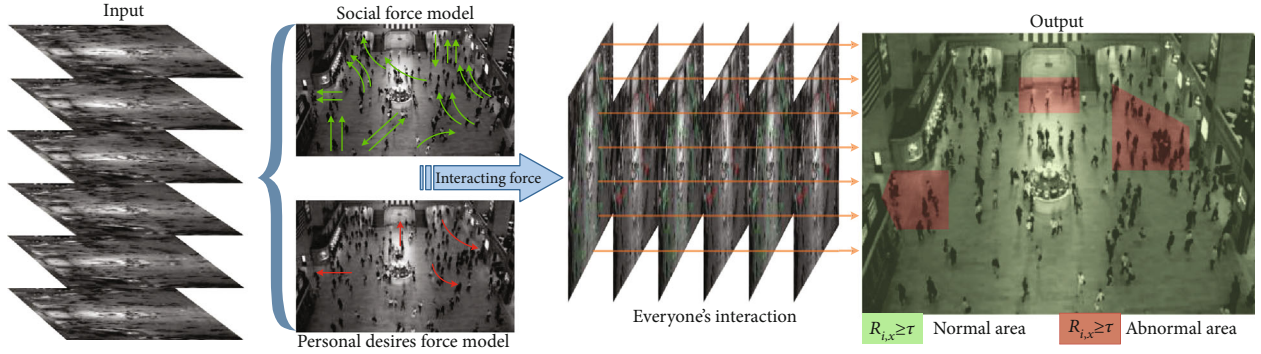


FIGURE 1: Detection of abnormal human behavior region. According to the interaction between social force and individual desire force, the behavior categories in video monitoring are marked out. Green stands for normal behavior and red stands for abnormal behavior.

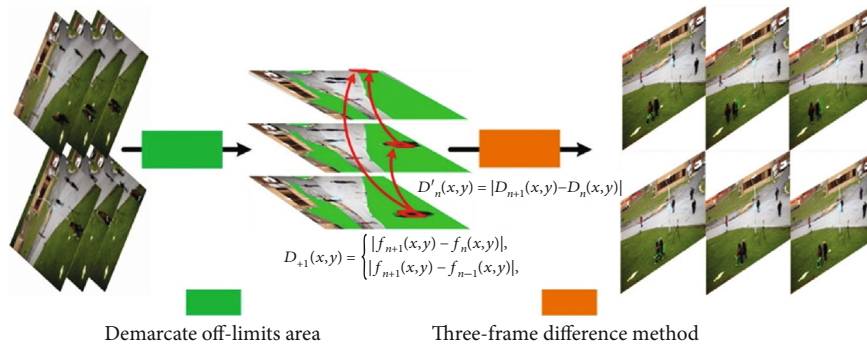


FIGURE 2: Intrusion detection model. Three-frame difference method is used to detect the intrusion.

analysis is needed to make an effective judgment on abnormal human behavior. The target, which is marked out by red in the previous step, will be tracked and monitored subsequently. In the process of target tracking, we put forward a multimodel based on the three-frame difference method to detect abnormal human behavior accurately.

3.2. Intrusion Detection. Intrusion refers to entering a designated area by a mobile target such as a person or an object without permission. In the monitoring area, the motion of the target is with a strong purpose. As a result, there are obvious changes between consecutive frames, and the position of the target is different in each frame. Interframe difference method is the difference between two or three consecutive frames of images in time. Pixel points corresponding to different frames are subtracted to derive the absolute value of gray difference. When the absolute value exceeds a certain threshold, it can be determined as a moving target, realizing the detection goal. The two-frame difference method is suitable for small and slow targets. When the target moves rapidly, the position difference between the two frames of the target image is too big, image ghosting might very well happen. Therefore, this paper applies the three-frame difference method to abnormal human behavior analysis. Let the images of frame $n-1$, frame n , and frame $n+1$ in the video sequence be f_{n-1} , f_n , and f_{n+1} , respectively. The corresponding gray value of pixel points should be $f_{n-1}(x,y)$, $f_n(x,y)$, and $f_{n+1}(x,y)$. Then, sub-

tract the gray value to obtain the difference in image D_{n+1} and D_n . In this paper, we optimize the three-frame difference method, that is, adding the difference between the third frame and the first frame:

$$\begin{aligned} D_{n+1}(x,y) &= |f_{n+1}(x,y) - f_{n-1}(x,y)|, \\ D_{n+1}(x,y) &= \begin{cases} |f_{n+1}(x,y) - f_n(x,y)|, \\ |f_{n+1}(x,y) - f_{n-1}(x,y)|, \end{cases} \quad (8) \\ D_n(x,y) &= |f_n(x,y) - f_{n-1}(x,y)|. \end{aligned}$$

Threshold processing and connectivity analysis are carried out on the difference image D_{n+1} and D_n to obtain the extracted target, as shown in Figure 2.

$$D'_n(x,y) = |D_{n+1}(x,y) - D_n(x,y)|. \quad (9)$$

3.3. Left-Behind Baggage Detection. Left-behind baggage refers to the situation that items such as bags or backpacks are left unattended in the monitoring area for more than a certain period. In order to detect the left-behind items more accurately, the Sobel operator is introduced into the three-frame difference method to detect the edges of the target. The purpose of this method is to determine the presence of left-behind items quickly and effectively when the item is left unattended. The Sobel operator monitors the horizontal edge in this paper. Compared with



FIGURE 3: Left-behind baggage detection. When the separation of human and objects is detected, the item would be marked out. When a pedestrian leaves the monitored area, the model outputs it as an abnormal behavior.

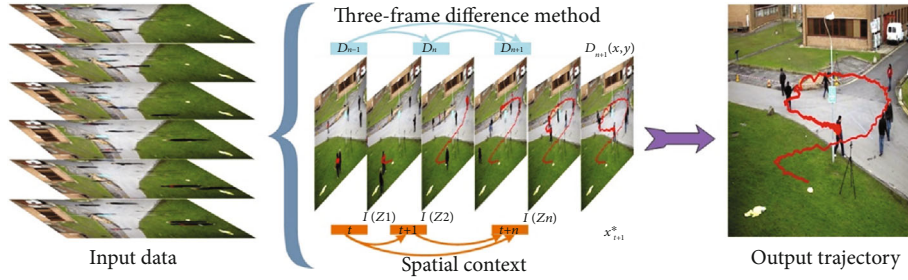


FIGURE 4: Motion trajectory identification. The spatial context model is established from the images obtained by the three-frame difference method. Human action often has a purpose. According to the movement or wandering route, this paper perceives it as abnormal behavior.

similar algorithms, the Sobel operator weights the pixel threshold, reducing the degree of edge blur effectively. Combined with the three-frame difference method, set the moving target as $R(x, y)$ and the left-behind object model as

$$r(x, y) = D'_n(x, y) - R(x, y). \quad (10)$$

In order to improve the reliability of the legacy warning, this paper adds the conditions for the disappearance of pedestrians to the remnant model. This method will provide an early warning when a pedestrian leaves the surveillance area. Therefore, the legacy warning model is updated to

$$r_T(x, y) = \begin{cases} \text{normal,} & \text{pedestrian exists,} \\ \text{abnormal,} & \text{pedestrian leaves.} \end{cases} \quad (11)$$

When the people and the target are separated, the pedestrian exists in the monitoring area and the model determines that it is a normal behavior. When the people and the target are separated, the pedestrian leaves the monitored area, and the model determines that it is an abnormal behavior and promptly makes an early warning, as shown in Figure 3.

3.4. Motion Trajectory Identification. With the improvement of video analysis, single or local image information cannot meet the precise analysis requirements. According to the characteristics that pixels in the space-time region are spatially smooth, this paper introduces the spatial context model [24] into the three-frame difference algorithm and proposes a new trajectory analysis model.

According to the three-frame difference method, $D'_n(x, y)$ is the target of monitoring. The confidence graph of each frame $C(x)$ in the video is obtained by the STC algorithm.

$$\begin{aligned} C(x) &= P(x|O) = \sum_{c(z) \in X^c} P(x, c(z)|O) \\ &= \sum_{c(z) \in X^c} P(x|c(z), O)P(c(z)|O). \end{aligned} \quad (12)$$

Based on the spatial context, the conditional probability $P(x|c(z), O)$ is the spatial relationship between the target and the residual scene. We define spatial context information as

$$X^c = \{c(z) = (I(z), z) | z \in \Omega_c(x^*)\}, \quad (13)$$

where $\Omega_c(x^*)$ is the remaining area excluding the target and $I(z)$ is the image feature of the target at the point z . The maximum confidence position of the confidence graph in the spatial region is the target position of the $t + 1$ frame.

$$x_{t+1}^* = \arg \max_{x \in \Omega_c(x_t^*)} C_{t+1}(x). \quad (14)$$

STC is introduced into the three-frame difference method. In the process of monitoring the target, the target position is tracked and predicted by the maximum confidence graph within the region, and the motion trajectory of the target is marked, as shown in Figure 4.

Based on the three-frame difference method, with the amalgamation of Sobel operator and spatial context relationship, a multimodel analysis method for abnormal human behavior is established. This method can accurately analyze and determine abnormal human behaviors such as intrusion,



FIGURE 5: Intrusion detection result.

left-behind baggage, motion trajectory, and make alerts in time.

4. Simulation Experiment

To verify the validity of this method, we select UMN, CAVIAR, and other datasets to test the algorithm.

Hardware configuration: in this paper, we use MATLAB under the Windows 10 system for experimental simulation. The simulation calculation runs on a small server with a CPU of E5-2630 v4, a main frequency of 2.2 GHz, and a memory of 32 GB.

Evaluation index: in order to verify the effectiveness of the method, both qualitative and quantitative criteria are utilized. In the qualitative assessment, we show the anomalous behavioral calibration of the dataset. In the quantitative assessment, we adopted the pixel-wise receiver of characteristics (ROC) and area under ROC (AUC) are employed. ROC representative of detection capabilities, while reflecting the relationship between the TPR and FPR. Among them, TPR is the true positive rate, $TPR = TP/P$, where TP is the true positive and P is the positive pixel number. FPR is false-positive rate, $FPR = FP/N$, where FP is the false positive and N is the negative pixel number.

When conducting intrusion detection in the monitoring area, we first mark the control area in red color. If there is any abnormality, the abnormal pedestrians will be marked

TABLE 1: Frame-level AUC comparison for intrusion detection in different datasets.

Data	Method (AUC %)		
	Literature [15]	Literature [17]	Ours
UMN	0.9325	0.9967	0.9681
CAVIAR	0.9173	0.9524	0.9597
USCD	0.9048	0.9250	0.9452
Average	0.9182	0.9580	0.9577

Note: boldface numbers indicate excellent indicators of each algorithm.

with green boxes, as shown in Figure 5. The monitoring objects of Video 1 and Video 3 are corridors. When students are in class or dismissed after school, there should be no person in the red area. If a target intrudes into the red area, the proposed method determines that this action is an illegal intrusion. The monitoring objects of Video 2 and Video 4 is public areas. Lawns, rooftops, and public areas are marked as red forbidden areas. If a target breaks into the red area, the proposed method will mark it out in time and give warnings. Video 5 belongs to the USCD dataset. In order to ensure the safety of the walking crowd, we set the other motion status indicators as abnormal. It can be seen from the simulation results that the algorithm effectively identifies the foreign object intrusion behavior in the video, such as bicycles, scooters, and minivans.

According to the experimental results, the proposed method is not only applicable to the indoor scene but also



FIGURE 6: Result of left-behind baggage detection.

to the outdoor scene. When the target enters the monitoring region, our method can identify the target quickly and effectively.

Table 1 shows the AUC comparison of different algorithms under different datasets.

In the UMN data set, the method proposed in literature [17] shows the best results. In the CAVIAR and USCD datasets, ours demonstrates the best results. Among the average of the three data sets, the method proposed in literature [17] shows excellent results, and ours is only 0.0003 different from the method.

When conducting a left-behind baggage experiment, we select surveillance data from parks, streets, and bus stop. This paper presents a method combining edge detection and the three-frame difference method, which can not only detect the state of separation between the owner and objects quickly but also determine the left-behind item accurately through the connection between each video frame. The detected item will be denoted with a red dot, as shown in Figure 6. When the pedestrian leaves the monitoring screen, the algorithm judges it as an abnormal behavior and issues a warning in time.

No matter where a person is going, his behavior must have a certain purpose. Therefore, the movement trajectory is regular, i.e., when people do something with a purpose, the movement trajectory tend to be a straight line. In contrast, aimless behavior will lead to a wandering state of the trajectory, as shown in Figure 7. The outputs of Videos 1, 2, and 3 are marked abnormal because the motion trajectories

present a wandering state, as opposed to trajectories in Video 4 and Video 5 which are close to straight lines, and marked normal.

In order to verify the generalization ability of the algorithm, we also selected the CVC-08 dataset and the JAAD data set for driving anomaly detection. The video dataset is captured by a car camera and the viewing angle is basically the same as the driver's daily driving. At the same time, in order to prove the adaptability of the algorithm to different targets such as people and cars, we keep the resolution of each frame of the video consistent with the above.

The results of the car-driving anomaly detection are shown in Figure 8. The CVC-08 dataset shows the pedestrian detection of the car during driving. The experimental results show that the algorithm can effectively identify when the pedestrian enters the zebra crossing and leaves the zebra crossing. The JAAD-01 video dataset is for bicycles crossing the road. The experimental results show that the algorithm can recognize the target when it is far away from surveillance. The JAAD-03 video dataset is a car replacement lane. When the car is changed from the rightmost lane to the left lane, the algorithm can effectively identify the vehicle under multiple targets. The JAAD-04 video dataset is for pedestrians crossing the road at night. The dataset demonstrates that the algorithm can still effectively identify the target in the absence of illumination.

Table 2 shows the AUC comparison of driving anomaly detection in CVC and JAAD. Both CVC-08 and JAAD-04 detect pedestrians crossing the road. In the normal weather

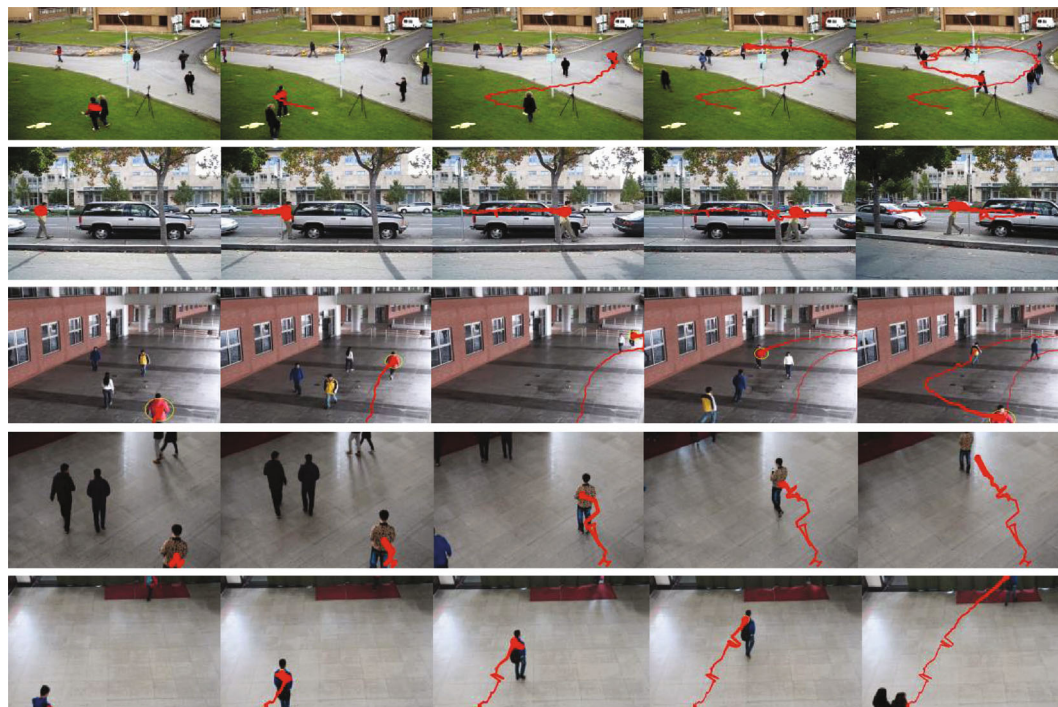


FIGURE 7: Result of motion trajectory identification.

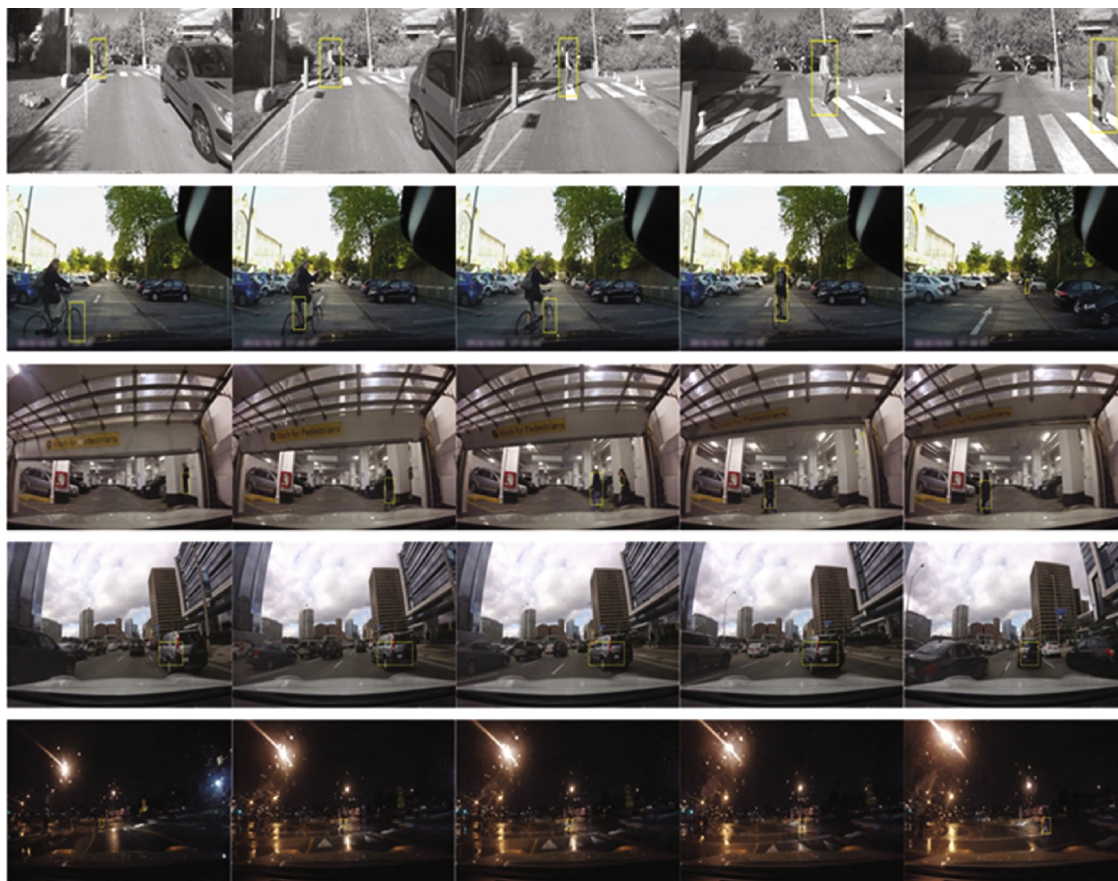


FIGURE 8: Result of driving anomaly detection.

TABLE 2: AUC comparison of driving anomaly detection in CVC and JAAD.

Data	Method (AUC %)		Ours
	Literature [16]	Literature [19]	
CVC-08	0.9063	0.8537	0.8893
JAAD-01	0.8817	0.8941	0.9051
JAAD-02	0.9085	0.8838	0.9166
JAAD-03	0.9203	0.9135	0.9087
JAAD-04	0.8739	0.8581	0.8932
Average	0.8981	0.8806	0.9026

Note: boldface numbers indicate excellent indicators of each algorithm.

conditions of CVC-08, the literature [16] proposed the method works best. In the case of poor JAAD-04 light, the method of this paper showed the best results. In the JAAD-02 video dataset, the method has the best recognition effect on bicycles. However, in the JAAD-04 video dataset, the method of this paper is not as good as the method proposed by literature [16]. Finally, in the JAAD-03 video dataset, this method demonstrates the applicability to multitarget.

5. Conclusions

This paper proposes a multistage analysis of abnormal human behavior in complex scenes to optimize the previous method, which relies on manual monitoring and is subject to time lag. Firstly, we differentiate the normal and abnormal areas of human behavior roughly in video surveillance according to the social force similarity measurement model. Secondly, based on the three-frame difference method, a multimodel which is suitable for intrusion and left-behind baggage detection, motion trajectory identification is established to do a precise analysis of abnormal human behavior. The experimental results show that the proposed method has good results in human abnormal behavior analysis and driving anomaly detection, which fully demonstrates the adaptability and generalization ability of the proposed method.

Data Availability

The dataset contains confidential information such as the performance parameters and tactical technical indicators of abnormal human behavior. Therefore, the dataset of this paper has certain confidentiality and cannot be released.

Conflicts of Interest

On behalf of my coauthors, the authors declare that there is no conflict of interests regarding the publication of this article.

Acknowledgments

This paper is supported by the National Natural Science Foundation of China (61703143), Science and Technology Department of Henan Province (192102310260), Scientific

and Technological Innovation Talents in Xinxiang (CXRC17004), the young backbone teacher training project of Henan University (2017GGJS123), and the Science and Technology Major Special Project of Xinxiang City (ZD18006).

References

- [1] X. Lan, S. Zhang, P. C. Yuen, and R. Chellappa, "Learning common and feature-specific patterns: a novel multiple-sparse-representation-based tracker," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2022–2037, 2018.
- [2] Y. Qi, S. Zhang, L. Qin et al., "Hedging deep features for visual tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 5, pp. 1116–1130, 2019.
- [3] S. Zhang, Y. Qi, F. Jiang, X. Lan, P. C. Yuen, and H. Zhou, "Point-to-set distance metric learning on deep representations for visual tracking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 1, pp. 187–198, 2018.
- [4] H. Zhang, J. Bai, Z. Li, Y. Liu, and K. Liu, "Scale invariant SURF detector and automatic clustering segmentation for infrared small targets detection," *Infrared Physics & Technology*, vol. 83, pp. 7–16, 2017.
- [5] M. ElMikaty and T. Stathaki, "Detection of cars in high-resolution aerial images of complex urban environments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 10, pp. 5913–5924, 2017.
- [6] H. Zheng, J. F. Liu, J. L. Gao, and Q. Lu, "Feature detection method for small targets of complex multimedia images in cloud environment," *Multimedia Tools and Applications*, vol. 76, no. 16, pp. 17095–17112, 2017.
- [7] U. Asif, M. Bennamoun, and F. Sohel, "Unsupervised segmentation of unknown objects in complex environments," *Autonomous Robots*, vol. 40, no. 5, pp. 805–829, 2016.
- [8] D. Luo and Y. Li, "Multi-stage and multi-attribute risk group decision-making method based on grey information," *Grey Systems: Theory and Application*, vol. 5, no. 2, pp. 222–233, 2015.
- [9] W. Matin, Z. Nasim, N. Ahmed, K. Zannat, and F. Muhammad, "Artificial neural network based system for intrusion detection using clustering on different feature selection," *International Journal of Computer Applications*, vol. 126, no. 12, pp. 21–28, 2015.
- [10] G. Bhartiya and A. S. Jalal, "Forgery detection using feature-clustering in recompressed JPEG images," *Multimedia Tools and Applications*, vol. 76, no. 20, pp. 20799–20814, 2017.
- [11] K. Huang, Y. Zhang, B. Lv, and Y. Shi, "Salient object detection based on background feature clustering," *Advances in Multimedia*, vol. 2017, Article ID 4183986, 9 pages, 2017.
- [12] A. Iosifidis, A. Tefas, and I. Pitas, "Class-specific reference discriminant analysis with application in human behavior analysis," *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 3, pp. 315–326, 2015.
- [13] B. Xiao, P. Georgiou, B. Baucom, and S. S. Narayanan, "Head motion modeling for human behavior analysis in dyadic interaction," *IEEE Transactions on Multimedia*, vol. 17, no. 7, pp. 1107–1119, 2015.
- [14] J. Pan, H. Hu, X. Liu, and Y. Hu, "Multiscale entropy analysis on human operating behavior," *Entropy*, vol. 18, no. 1, p. 3, 2016.

- [15] S. Yi, H. Li, and X. Wang, "Pedestrian behavior modeling from stationary crowds with applications to intelligent surveillance," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4354–4368, 2016.
- [16] J. Hu, L. Xu, X. He, and W. Meng, "Abnormal driving detection based on normalized driving behavior," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 8, pp. 6645–6652, 2017.
- [17] Y. Yuan, J. Fang, and Q. Wang, "Online anomaly detection in crowd scenes via structure analysis," *IEEE Transactions on Cybernetics*, vol. 45, no. 3, pp. 548–561, 2015.
- [18] Y. Yuan, D. Wang, and Q. Wang, "Anomaly detection in traffic scenes via spatial-aware motion reconstruction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 5, pp. 1198–1209, 2017.
- [19] Y. Yu, L. Cao, E. A. Rundensteiner, and Q. Wang, "Outlier detection over massive-scale trajectory streams," *ACM Transactions on Database Systems*, vol. 42, no. 2, pp. 1–33, 2017.
- [20] W. Nie, X. Li, A. Liu, and Y. Su, "3D object retrieval based on spatial+LDA model," *Multimedia Tools and Applications*, vol. 76, no. 3, pp. 4091–4104, 2017.
- [21] B. Leng, J. Zeng, M. Yao, and Z. Xiong, "3D object retrieval with multitopic model combining relevance feedback and LDA model," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 94–105, 2015.
- [22] L. Chen, X. Hu, T. Xu, H. Kuang, and Q. Li, "Turn signal detection during nighttime by CNN detector and perceptual hashing tracking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3303–3314, 2017.
- [23] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 935–942, Miami, FL, USA, 2009.
- [24] L. Bai and Q. Chen, "Visual phrase recognition by modeling 3D spatial context of multiple objects," *Neurocomputing*, vol. 253, pp. 183–192, 2017.



Hindawi

Submit your manuscripts at
www.hindawi.com

