



Recognition of Human Abnormal Behavior in Static Image of Intelligent Monitoring System Based on Neural Network Algorithm

Hai-jing Zhou^(✉)

Chongqing Vocational Institute of Tourism, Chongqing 409099, China
zhouhaijing554@yeah.net

Abstract. In recent years, human abnormal behavior recognition in static images is a hot topic in the vision field of intelligent monitoring system. In order to achieve the goal of accurate recognition of human abnormal behavior in static images, a method of human abnormal behavior recognition in static images of intelligent monitoring system based on neural network algorithm is proposed. The pedestrian detection algorithm based on neural network algorithm for filtering channel features is used to locate each target human body, collect feature data and denoise. Then the appearance model of each human structure feature frame is constructed, and finally the neural network algorithm is used for behavior classification and recognition. The experimental results show that the human abnormal behavior recognition method based on neural network algorithm in the static image of intelligent monitoring system can recognize the behavior of multiple human bodies in a single frame image. This method can provide multi category classification of human abnormal behavior at the same time, the experimental effect is obvious, and the accuracy is high.

Keywords: Neural Network · Intelligent monitoring · Static image · Behavior recognition

1 Introduction

Human abnormal behavior recognition in static images has a wide range of application prospects and potential economic value in security monitoring, advanced human-computer interaction, video conference, behavior based video retrieval and medical diagnosis. It is a research hotspot in the field of computer vision. Video intelligent monitoring system is one of its most important applications [1]. At present, the detection, tracking and recognition of static target in intelligent monitoring system is still in the primary stage, and the behavior analysis of static target is relatively simple. However, due to the interference factors, the recognition effect is still difficult to meet the current research objectives. Based on this, this paper discusses the common problems of human abnormal behavior recognition in the monitoring scene, and defines a small amount of behavior on the scene as abnormal behavior. Based on this, this paper studies the human abnormal

behavior recognition method in video surveillance, and expounds the research significance, development status and application fields of human abnormal behavior detection and recognition method [2]. In the field of computer vision, crowd abnormal behavior detection technology can be widely used in video surveillance, intelligent video analysis, group behavior recognition and other fields. Jiang Jun et al. [3] proposed a crowd abnormal behavior detection method based on pulse line flow convolutional neural network using deep residual network as the backbone network, extracting the apparent characteristics of crowd behavior through the input of the spatial network, and using the pulse line Stream extracts the movement characteristics of crowd behaviors to further improve the accuracy of crowd abnormal behavior detection. Obtain significant spatial characteristics of crowd behavior through convolutional neural network, and obtain temporal characteristics of crowd behavior through pulse flow combined with convolutional neural network; finally, the output of the two networks is averaged to complete the detection of abnormal crowd behavior. In the process of criminal investigation, Wu Songwei [4] and others used visual-based behavior analysis and understanding technology to discover and identify the abnormal behavior of potential perpetrators in a timely manner, which has extremely important practical significance for crime early warning. By designing the hardware and software system framework of the video surveillance abnormal behavior detection system, researching video-based pedestrian detection methods and abnormal behavior detection and recognition methods, and completing the implementation and development of related algorithms on the system, thereby improving the accuracy of the system and efficiently detecting abnormal behaviors ability. Zhang Xin and Qi Hua [5] studied and analyzed the complex and diverse problems of abnormal human behavior characteristics in view of the situation where the classification of human behavior in the surveillance video in public places is greatly affected by the background. Based on the YOLOv4 network model, a network human body abnormal behavior detection model is established. The experimental data results are obtained through K-Means clustering, and the processed data samples are sent to the three models for training. The test results show that it can meet the requirements of high accuracy and real-time. The advantages and disadvantages of the existing methods are summarized and analyzed, and some key technical methods are studied and improved. The main contents of this paper are static target detection, static target tracking, human behavior feature extraction and human behavior recognition methods. In static target detection, the parameters of hybrid Gaussian model are initialized by neural network method, which avoids the influence of static target on the initialization of Gaussian model [6]. The recognition method based on neural network can adaptively select the number of recognition points, and improve the accuracy and efficiency of static target recognition.

2 Recognition of Human Abnormal Behavior in Static Image of Intelligent Monitoring System

2.1 Collection of Abnormal Behavior Characteristics of Human Body in Static Image

Aiming at the problem of human abnormal behavior recognition in monitoring scene, a method of human abnormal behavior recognition based on neural network is proposed.

Establish the relationship between human static characteristics and behavior cognition, describe the middle characteristic behavior of human behavior, and realize the characteristic discrimination of human abnormal behavior [7]. In order to ensure the effect of feature recognition, firstly, the feature parameters of human abnormal behavior in static images are collected, and the acquisition steps are optimized, as shown in Fig. 1

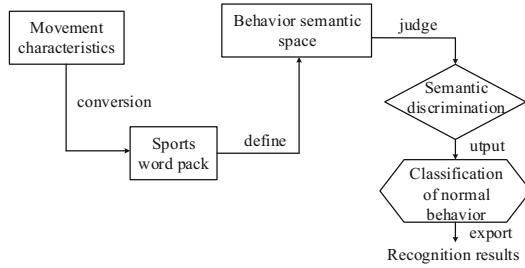


Fig. 1. Human behavior characteristics collection process of static image

In the process of feature recognition, it is necessary to use time-space interest points as low-level static features. The same kind of human behavior has similar concept distribution. The static distribution structure can be established to divide the human behavior categories and judge the specific behavior feature categories [8]. In the process of feature collection, in order to ensure the invariance of behavior features, better reflect the static information, and realize the comprehensive complementary of intelligent monitoring system features and global features, it is the key to realize the recognition of human abnormal behavior in static images. Therefore, it is necessary to further optimize the feature data processing steps based on neural network theory [9]. The platform of human behavior action capture is composed of a reference module of attitude heading and a computer, which can capture the body, upper and lower arms and legs. When the platform starts to work, the attitude and heading reference module uploads the attitude data of human action capture at 30 Hz frequency through the wireless network. The data is received by the computer, and the data sent by the attitude and heading reference module is received by the address query mode [6]. In order to reduce the volume and electromagnetic interference, it is necessary to enhance the platform independence and use the extended neural network for real-time data acquisition. After de-noising processing, data fusion is carried out to capture attitude angle, angular velocity and acceleration in real time. Then, the data is sent to the upper computer through the wireless network module, and the data is processed by the upper computer program, which is transformed into the signal form [7].

The neural network is used to collect the original data, and the relative attitude angle is obtained through the conversion relationship between Euler angle and quaternion [8]. The absolute attitude angle can be calculated by collecting the acceleration and using the direction cosine transformation of the magnetic field between the geographic coordinate system and the neural network coordinate system. When using the neural network to calculate the attitude, the data is divergent and cannot be used for a long time. However,

the accelerometer is vulnerable to vibration and its dynamic performance is poor. Therefore, it is necessary to fuse the data generated by the accelerometer to obtain a stable and accurate attitude angle. According to the vibration intensity, the accelerometer filter gain is updated in real time to ensure the dynamic response characteristics of the attitude and heading reference module and eliminate the steady-state error quickly. According to the three-axis acceleration value processed by neural network, the neural network is used to calculate the tilt angle roll and pitch angle pitch, and the yaw angle yaw is calculated according to the tilt compensation magnetometer value. According to the state vector, the high-precision attitude angle is calculated by using Euler angle conversion matrix. Through the tilt angle can capture human walking, running behavior signal, pitch angle can capture human static behavior signal, yaw angle can capture human jumping, squatting behavior signal. In fact, the complexity of the collected images is much greater than that of the data set used in this paper, including the influence of weather, light, angle of view and occlusion. It is a challenging task to mine practical multi person abnormal behavior recognition methods, which can better reduce the complexity of the algorithm, improve the running speed and so on. The processing of human abnormal behavior features in static images of intelligent monitoring system generally includes the following processes: static target detection, target tracking, feature selection and static representation, behavior recognition and understanding. Figure 2 is the whole process of human abnormal behavior processing in the static image of an intelligent monitoring system.

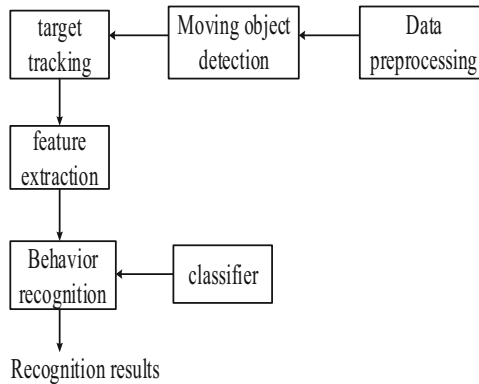


Fig. 2. Steps of human abnormal behavior processing in static image

In the process of human abnormal behavior recognition in static images, there is diversity due to the change of video content. It is necessary to capture the video behavior features based on the collected results and abstract them into data to provide further description of intermediate behavior features for video description [10]. Because in the process of human abnormal behavior recognition in static images, the recognition content of human behavior features is combined in an unknowable way, so it is necessary to use neural network for human behavior derivation and recognition [11]. Aiming at the problem of human complex behavior modeling with a long time span and a variety of basic actions, a collaborative clustering algorithm is used to model the correlation

between static words and human actions synchronously. At the same time, neural network is used to model the temporal correlation of human behavior, so as to improve the recognition accuracy of human complex behavior.

2.2 Optimization of Human Abnormal Behavior Recognition Algorithm in Static Image

Static object detection is to extract the change region from the background in the video image sequence, which is the basic and key part of surveillance video analysis, because the quality of static object segmentation directly affects the subsequent processing. In the process of image recognition, the image background will be interfered by the external environment, so the static target extraction becomes a more difficult work [12]. In practical applications, in order to better adapt to the static target detection algorithm of various environmental changes, combined with inter frame difference method, optical flow method, background subtraction method for denoising. By separating the changing region from the background image, the static target region and the background region can be segmented correctly. Multi path recognition of human motion features mainly involves image feature analysis, processing, feature extraction and pattern recognition. The general recognition process is shown in Fig. 3.

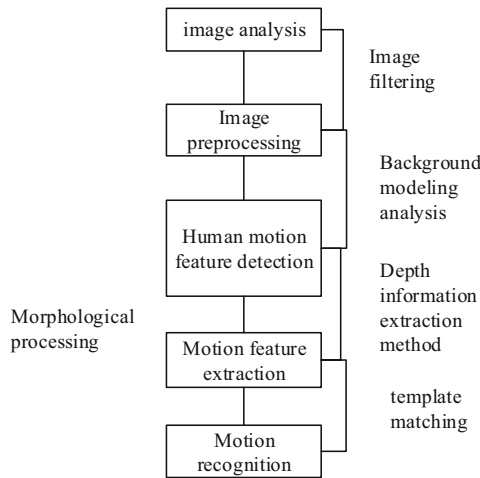


Fig. 3. General recognition process of human body features

Before motion feature detection, it is necessary to preprocess the acquired image information to filter the noise and provide accurate data for subsequent recognition. The human motion feature detection step is set to extract the related feature points from the complex human motion region. Under the occlusion condition, the clear human image can be obtained through morphological processing. In order to judge which type of motion the current motion belongs to, the appropriate human body representation method is selected and the useful motion information is extracted from the human body

sequence. Starting from the general recognition process, the multi-path recognition of human motion characteristics is carried out by using depth information.

In the process of abnormal behavior recognition, we need to introduce weighting coefficient to segment background and static target effectively. Chaotic particle swarm optimization is used to improve the iterative speed of image processing. Based on the commonly used description of posture and static features, the static features of the edge and limb shape of the video image are collected and judged to achieve de-noising [13, 14]. Combined with the neural network algorithm, the time-space feature of the image is detected, and the $I(x, y)$ structure coordinate of the image to be recognized is transformed into the image feature sequence $I(x, y, t)$. In order to detect features from image sequence $I(x, y, t)$, a separable linear filter is applied to construct the following response function R for a single fixed camera video

$$R = (I(x, y, t) \times h_{ev})^2 + (I \times g(x, y, \sigma) \times h_{od})^2 \tag{1}$$

Among formula (1):

$$g(x, y; \sigma) = \frac{e^{(-(x^2+y^2)/2\sigma_l^2 - t^2/2\tau_l^2)}}{\sqrt{(2\pi)^3 \sigma_l^4 \tau_l^2}} - 1 \tag{2}$$

Then the smoothing kernel of the static image is as follows:

$$h_{ev}(t; \tau, \omega) = -R \cos(2\pi t\omega) e^{t^2/\tau^2} \tag{3}$$

$$h_{od}(t; \tau, \omega) = R - \sin(2\pi t\omega) e^{t^2/\tau^2} \tag{4}$$

If the feature weight of the static image is ω , e^{t^2/τ^2} is the loss vector under the t-interference environment. τ is a still image smoothing coefficient, which has a great influence on the mesh performance. When the change of τ value is small, the error rate of human abnormal behavior recognition in static image is almost constant. Based on this, in the process of human abnormal behavior recognition in static images, when the sample data s is input as the original feature data of the model, the deviation value is B . Furthermore, the activation function is introduced based on the nonlinear calculation principle, and the normalized sample C is obtained by reducing the dimension of the merging layer.

$$C = Bsh_{od}(t; \tau, \omega) \cdot Xh_{ev}(t; \tau, \omega) \tag{5}$$

Assuming that the feature image sample to be recognized is p , the sample matrix is regularized. And record it in the feature image D . Then the standard sample vector and the standard learning sample vector are calculated

$$E = p \sum_1^k |D - CK|^2 \tag{6}$$

In formula (6), E represents the Euclidean distance between the sample to be classified and the learning sample, and the number of all kinds of samples is the same, which

is k . The Gaussian function of smoothing factor $E = 0.5$ is selected as the activation function to obtain the initial probability formula.

$$\Delta P = e^{-D/2a^2} / 2VE \tag{7}$$

In the above algorithm, M is a video sequence with m , and the sample set is V , and it is recorded as $V = \{v_1, v_2, \dots, v_M\}$. Using the difference between the current two frames in the video image sequence or between the current frame and a previous frame to extract the static region in the image for denoising. The principle is shown in Fig. 4.

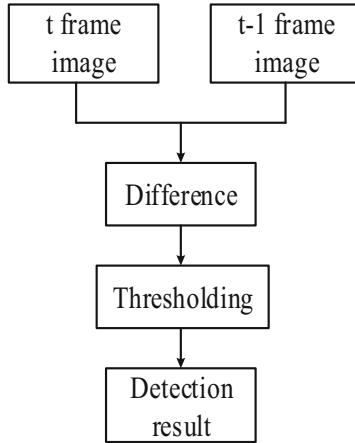


Fig. 4. Denoising of characteristic parameters of inter frame difference

Inter frame difference is the difference operation of the corresponding pixels between two frames. When the illumination condition in the scene is relatively stable, the difference value of the background part in the image sequence is zero. There is noise in the actual scene, and the difference between the regions of non-static objects may not be zero, which will affect the detection results of static objects. Based on this, we need to thresholding the difference results. By thresholding the detection results, the binary image with only static target is obtained, which can effectively suppress the noise interference. The background model is used to update the background image, and then the current frame and the background image are used to get the static target detection technology. The background subtraction method is also a special inter frame difference method. This method can get a more complete static target with fast operation speed, and can adapt to the real-time requirements of the system. When using this method to detect static targets, it is very important to get a more stable background model. Based on this, the background subtraction method is further optimized, and the specific steps are shown in Fig. 5.

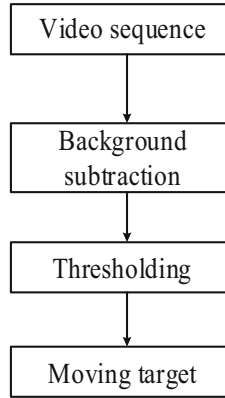


Fig. 5. Background subtraction processing steps

The background subtraction method is used to directly compare the gray value of the background image and the pixel of the static object entering the background image. Once the difference between the gray value of the pixel of the two images is greater than a set threshold, it indicates that there is an abnormal phenomenon. The mathematical formula can be expressed as follows:

$$F(x, y) = \begin{cases} 1, & |f_1(x, y) - f_2(x, y)| > S \\ 0 & \text{other} \end{cases} \quad (8)$$

In formula (8), $f_1(x, y)$ is the background image, $f_2(x, y)$ is the background image when the abnormal phenomenon occurs, $F(x, y)$ is the binary image on (x, y) , and S is the initial threshold. A window with 512 sample points is used to segment the original signal, and its rectangular window is used to segment the original signal. After windowing, a single signal contains 1024 samples. If the sampling frequency of the signal is 100 Hz, the time span of the signal after windowing is 10.24 s, which is enough to complete a single action. The acceleration signal is processed by feature extraction to obtain an original feature matrix, which is trained to recognize the behavior signal.

The windowing method is used to process the processed human behavior signal. The sensor signal is extracted from the overlapped half window factory rectangular window. After the window factory is appropriate, a variety of features are extracted from a single window factory signal to form a feature vector to represent human behavior. Standard deviation is one of the statistical characteristics, which effectively reflects the degree of data dispersion of sensors. Because people are in static state, acceleration data are basically unchanged, and the standard value is 0. The acceleration data of people are constantly changing during the course of motion, and the standard value is greater than 0. The specific formula is:

$$a = \sqrt{\frac{1}{SF(x, y)} \sum_{i=1}^S (A_i - \bar{A})^2} \quad (9)$$

In formula (9), S is the number of samples and \bar{A} is the average value of samples. According to the formula, the static and dynamic behaviors of human body can be distinguished. Skewness can measure the skew direction of acceleration sensor data distribution, and analyze the gravity direction diagram. The specific formula (10) is as follows.

$$P = \frac{S \sum_{i=1}^s (A_i - \bar{A})^3}{(S - 1)(S - 2)a^3} \tag{10}$$

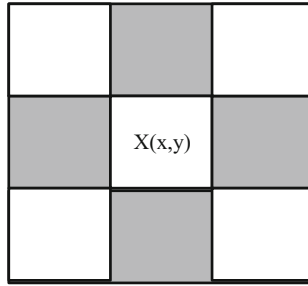
According to formula (10), skewness can effectively distinguish the human body jumping and squatting behavior. Kurtosis change curve can directly reflect the change degree of all signals at the peak of data curve, which is an important statistical feature. The specific calculation formula (11) is as follows.

$$Q = \frac{\sum_{i=1}^s (A_i - \bar{A})^4 n_i}{Sa^4} \tag{11}$$

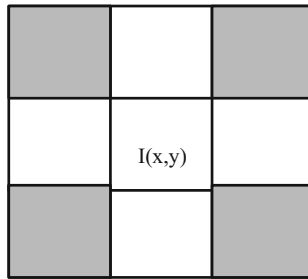
In formula (11), n_i is the sample interval. The formula can effectively distinguish human running behavior. The correlation coefficient can be used to measure the degree of linear correlation between variables. The specific calculation formula (12) is as follows.

$$W_{AB} = \frac{\sum_{i=1}^s (A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^s (A_i - \bar{A})^2} \sqrt{\sum_{i=1}^s (B_i - \bar{B})^2}} \tag{12}$$

In formula (12), A and B represent correlation coefficient variables respectively. According to the formula, combined with the correlation coefficient of gravity YZ, we can effectively distinguish between human walking and going upstairs. When the static target stays in a certain position for a long time, the pixels of the static target have a higher probability of occurrence in that position. Based on this, in the process of neural network updating, the target pixels will get a larger weight, so the target pixels will be misjudged as background pixels, and then the static target will be extracted. In this way, the static target will be missed and the hole will be formed. So the extraction of static target is not complete. Therefore, we need to use a morphological processing method to deal with the whole static target. The common algorithms in image processing include expansion, corrosion, open operation and close operation. The language of mathematical morphology is set theory, which uses special structural elements to process images. Common structural elements include 4-point neighborhood and 8-point neighborhood of 3×3 . The definition of 3×3 neighborhood is shown in Fig. 6.



(a) 4-domain



(b) \$field

Fig. 6. Definition of neighborhood background features of surveillance video

The main reasons for the change are the entry or exit of the monitoring target in the monitoring scene, the change of light and the influence of noise. In order to solve this problem, the initial threshold S is used as the threshold to distinguish the segmented object from the actual background. Set the pixel greater than or equal to the threshold as the object, and the pixel less than the threshold as the background. Based on this method, the feature data of static image can be extracted effectively, and the interference error in the process of image recognition can be eliminated effectively.

2.3 Realization of Abnormal Recognition of Static Portrait in Monitoring System

Static target detection and tracking is the research foundation of target behavior analysis and recognition in intelligent video surveillance system. Static object detection and tracking is to extract the static object in video sequence by image processing technology. Then, according to the characteristics of the static target, the static trajectory of the target is obtained. Finally, the behavior of the target can be analyzed and recognized according to the trajectory change and static characteristics of the target. Static object detection is the bottom module of computer vision, which aims to separate the static object from the background image. The effect of static detection will directly affect the accuracy of the following target tracking and behavior analysis. According to the current situation, it is very important to adopt the appropriate static target detection algorithm. The new video frame X_t and the B Gaussian distribution of the background are matched one by one. If

X_i matches one of the B Gaussian distributions, it means that the pixel is the background, otherwise it is the foreground. Generally, the gray value of background pixel is assigned as 0, and the gray value of foreground pixel is assigned as 1. In this way, the static object can be extracted from the background. Based on this, the algorithm flow structure of neural network is optimized, as shown in Fig. 7.

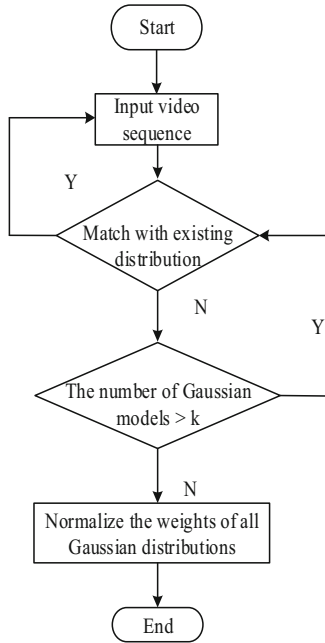


Fig. 7. Flow chart of portrait monitoring algorithm based on Neural Network

In the neural network, it is assumed that the probability of background pixels is high. Therefore, the Gaussian distribution with the highest probability of a pixel is used to represent the background model of the pixel. Human static object tracking is to find the position of the static human body in the current frame according to the position of the human body in the previous frame, and match the static human body in the two frames. If it can match, it means that the static human body in the two frames belongs to the same person. At present, the main reason for the slow progress of research on multi person abnormal behavior detection system is the limitation of universality and robustness. It is very difficult to obtain the behavior characteristics of multiple people and identify and classify them into complex scenes. The recognition method of human abnormal behavior in static image of intelligent monitoring system based on neural network algorithm mainly improves the speed and accuracy of the system from three aspects. The filter channel feature detector is used to calibrate the position of all human bodies in the image. It can reduce the influence of background on pose estimation and determine the position of some parts of the human body quickly and accurately. Through the neural network algorithm to obtain human static features and classification,

the accuracy of classification is guaranteed and the running speed is improved. In order to ensure the recognition effect of human abnormal behavior, region tracking method and active contour tracking method are used to recognize human abnormal static feature multi-path. The specific identification process is shown in Fig. 8.

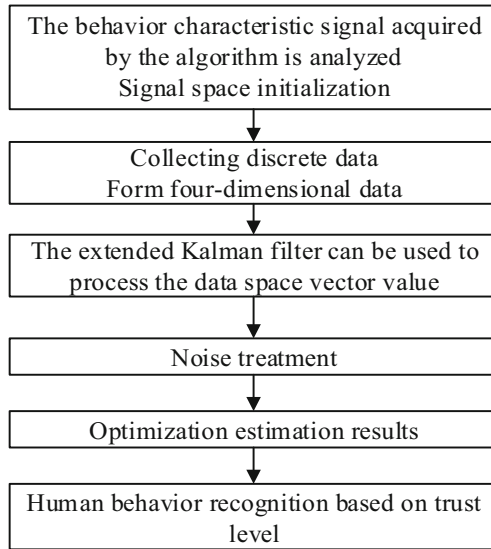


Fig. 8. Specific recognition process of portrait features

According to the identification process, the trust level of each component of the data space can be set. The optimal characteristic data is obtained by drying each grade. So as to better protect the effect of abnormal behavior recognition and improve the accuracy of human abnormal behavior recognition.

3 Analysis of Experimental Results

In order to verify the practical application effect of human abnormal behavior recognition in the static image of intelligent monitoring system based on neural network algorithm, the experimental detection is carried out. The experimental process is generated by the model and represented by behavior collection and display. Finally, the method described in the previous section is used to detect abnormal behaviors online and classify normal behaviors in real time. In this experiment, we use matlabr2019a platform to detect static objects in a video with two static human bodies in the behavior database of the Institute of automation, Chinese Academy of Sciences. Using 2.0 GHz 7 64 bit processor, 8 GB ram random access memory, Dx11 model graphics adapter. The experiment uses a host computer with the parameter configuration shown in Table 1.

Table 1. Experimental host configuration

Parameter	Index
built-in	USB3.0 bus
Graphics card	NAIDIA
processor	64 bit
CPU	IntelCore

Under the condition of Microsoft Kinect camera, the depth sensor is used to establish the experimental database. The specific data are shown in Table 2.

Table 2. Experimental data

Database	Resolving power	Remarks
MSR	320 * 2240	15fps; TXT file to the clearance node
RGBD	320 * 240	Indoor daily exercise; single person; whole body
LIRIS database	640 * 480	Indoor office action or interaction; single person or whole body
Chalearndatabase	640 * 480	16 daily behaviors; 15fps; TXT file joint points
Kitchendatabase	640 * 480	Camera installation; sports oriented
ReadingAct database	320 * 240	19 movements; 30 FPS; two cameras

In order to verify the practical application effect of this method, experimental detection is carried out. In order to ensure the accuracy of the experiment, the evaluation index of each behavior is standardized, and the detection accuracy, accuracy and recall rate of the four abnormal behaviors under this method are given. The specific results are shown in Table 3.

Table 3. Evaluation index of experimental results

This paper	Accuracy	Recall	Accuracy
Slap	0.900	0.888	0.932
Robbery	0.903	0.932	0.821
Shooting	0.903	0.865	0.923
Boxing	0.909	0.881	0.912

It can be seen from Table 3 that the method in this paper has a good ability to recognize abnormal human behaviors in four different types of static images. In terms of

abnormal feature recognition, the average recognition rate of the system in this paper is 90%, the image recall rate for abnormal behaviors is 87%, and the overall average effect of the system is 89%. This shows that the system in this paper has a good recognition rate for abnormal behavior detection. The result of human behavior is conducive to the stability of the detection system, so the static behavior data of the human body is collected according to the time step. Through the data measurement and analysis of the standard deviation, skewness, peak value and correlation coefficient of the X, Y, and Z axes. The results of the study are shown in Fig. 9.

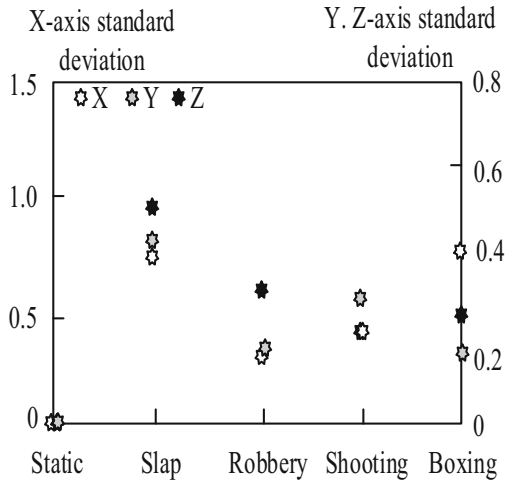


Fig. 9. Human behavior capture results

When human behavior is at rest, the standard deviation of X, y and Z axes is 0. When the human body is walking, the standard deviations of X, y and Z axes are 0.7, 0.45 and 0.5 respectively. When the human body is running, the standard deviations of X, y and Z axes are 0.3, 0.22 and 0.35 respectively. The standard deviations of X, y and Z axes are 0.4, 0.3 and 0.25 respectively when the human body is in the state of jumping up. When the human body is squatting, the standard deviations of X, y and Z axes are 0.8, 0.2 and 0.3 respectively. When the human body is in a static state, the component interval value does not change, so the study of human static abnormal behavior. By comparing the value range of various human behaviors, it can be seen that the traditional recognition method will mistakenly recognize the running state after walking faster. When the running speed is increased, it will be mistakenly identified as jumping state, if the jumping range is small, it will be mistakenly identified as squatting state, and the experimental results are inconsistent with the behavior habits.

The selected detector based on channel filtering features can effectively and accurately mark the position of each human body in the image, which can provide a strong guarantee for the subsequent pose estimation and improve the success rate of behavior recognition. In order to show the practical application effect of this method more clearly, the experiment was carried out by comparing it with the traditional recognition method, and 2–10 people’s behavior images were randomly identified and tested. The experimental time is controlled in the range of 928 s. On this basis, a static real image is randomly selected to detect the recognition results. The specific detection results are shown in Fig. 10.

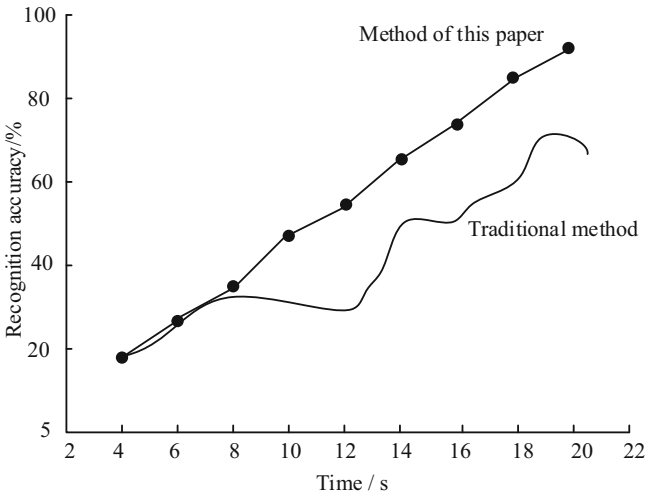


Fig. 10. Comparison test results

According to the test results in Fig. 10, the comparative test results of the traditional method and the method in this paper can be obtained. We can see that both the traditional method and the method in this paper have good recognition results in the range of less than 10 s, but as time continues to increase, the traditional recognition ability has eased, and the recognition accuracy is maintained at about 60%. However, with the increase of time in this paper, the recognition accuracy reaches 98%. It can be concluded that, compared with the traditional method, the human abnormal behavior recognition method based on neural network algorithm in this paper has a higher conversion rate in the actual application process, and can better meet the static human abnormal recognition target. In addition, the delays of abnormal behavior recognition under different methods are compared and recorded. The specific test results are shown in Table 4.

Table 4. Contrast detection of recognition delay between the two systems

Traditional recognition system				This paper introduces the recognition system			
Image acquisition	dynamic tracking	Dynamic identification	Accuracy%	Image acquisition	dynamic tracking	Dynamic identification	Accuracy%
—	S	S	59	—	S	S	69
S	—	S	86	S	—	S	94
S	S	—	70	S	S	—	91
—	—	S	55	—	—	S	79
—	S	—	60	—	S	—	71
S	—	—	40	S	—	—	52
S	S	S	100	S	S	S	100

From Table 4, we can get the image acquisition effect, image tracking ability, the traditional method under the dynamic recognition effect and the image recognition delay of the method in this paper. S represents the normal state of recognition. Through comparative analysis, it is found that there is a delay in the image recognition process. The image recognition delay rate of traditional methods is about 70% on average. In the static image of the intelligent surveillance system proposed in this paper, the neural network algorithm is used to recognize abnormal human behaviors. The method has high accuracy in the actual application process, the recognition accuracy rate is 89%, and the delay performance is 90%. Based on the above experimental results, the overall performance of the system algorithm in this paper is good, and the running speed has also been improved to a certain extent, which can better meet the requirements of abnormal behavior research on static human images.

4 Conclusion

The neural network algorithm can play a guiding role in the recognition of abnormal human behaviors in static images, and can recognize the behaviors of multiple humans on a single frame of images, and provide a multi-category classification method of abnormal human behaviors. Automatically learn human behavior models through video training sets without any manual annotations, and can apply the learned human behavior models to detect abnormal behaviors and recognize normal behaviors in real time. This paper builds the appearance model of the human structure feature framework, collects the abnormal behavior characteristics of the human body in static images, performs denoising processing according to the characteristic data, and uses the chaotic particle swarm search method to improve the iterative speed of image processing. Use neural network algorithm features to analyze the trajectory changes and static characteristics

of the target object, analyze and recognize the behavior of the target, and optimize and upgrade the recognition effect of abnormal human behavior in the static image according to the actual parameter results obtained, and improve the system design. This research is based on the recognition of abnormal human behavior in static images of the intelligent monitoring system based on neural network algorithm, which overcomes the shortcomings of accuracy and efficiency of traditional methods in the process of modeling complex human behavior. It alleviates the behavior type ambiguity caused by the lack of sufficient visual evidence in the real-time behavior recognition process of the system, so that the system can better complete the tasks of real-time abnormal behavior detection in the monitoring scene and online normal behavior recognition. It is conducive to optimizing the system update and upgrading, effectively improving the response speed of the system to recognize human behavior, reducing system memory usage, and improving the overall recognition effect of the system. In the actual application process of the system in this paper, the detection and recognition rate is 89%, the accuracy is 90%, and the delay is low, which shows that the method in this paper has higher accuracy, better stability, and more obvious effect.

Fund Projects. 1. The 2019 Municipal Education Commission Science and Technology Research Program Project, Project Name: Research on Attitude Adaptive Static Image Human Behavior Recognition(Project Number: KJQN201904601)

2. School level project of Chongqing Vocational Institute of Tourism in 2020: Project Name: Research On Human Action Recognition Algorithm Based On Deep Learning (Project Number:XJKT-2020-12)

References

1. Gorban, A.N., Mirkes, E.M., Tugin, I.Y.: How deep should be the depth of convolutional neural networks: a backyard dog case study. *Cogn. Comput.* **12**(1), 388–397 (2020)
2. Chen, S., Shen, Y., Yan, Y., et al.: Cholesky decomposition based metric learning for video-based human action recognition. *IEEE Access* (99), 1 (2020)
3. Jun, J., Zhuojun, Z., Mingliang, G., et al.: An abnormal crowd behavior detection method based on streak flow CNN. *Adv. Eng. Sci.* **52**(6), 215–222 (2020)
4. Songwei, W., Jun, L., Changjun, F., et al.: Design and implementation of video-based abnormal behavior detection system for criminal investigation. *Comput. Era* (9), 67–71,75 (2020)
5. Xin, Z., Hua, Q.: Research on human abnormal behavior detection algorithm based on YOLOv4. *Comput. Digit. Eng.* **49**(4), 791–796 (2021)
6. Al-Obaidi, S., Al-Khafaji, H., Abhayaratne, C.: Modeling temporal visual salience for human action recognition enabled visual anonymity preservation. *IEEE Access* **8**(9), 213806–213824 (2020)
7. Dhiman, C., Vishwakarma, D.K.: View-invariant deep architecture for human action recognition using two-stream motion and shape temporal dynamics. *IEEE Trans. Image Process.* (99), 1 (2020)
8. Wang, Z., Li, N., Wu, T., et al.: Simulation of human ear recognition sound direction based on convolutional neural network. *J. Intell. Syst.* **30**(1), 209–223 (2020)
9. Veni, S., Anand, R., Mohan, D., et al.: Feature fusion in multimodal emotion recognition system for enhancement of human-machine interaction. *IOP Conf. Ser. Mater. Sci. Eng.* **1084**(1), 012004 (10p) (2021)

10. Liu, S., Bai, W., Liu, G., et al.: Parallel fractal compression method for big video data. *Complexity* **2018**, 2016976 (2018)
11. Liu, S., Fu, W., He, L., Zhou, J., Ma, M.: Distribution of primary additional errors in fractal encoding method. *Multimedia Tools Appl.* **76**(4), 5787–5802 (2014). <https://doi.org/10.1007/s11042-014-2408-1>
12. Liu, S., Ma, X., Wu, H., et al.: An end to end framework with adaptive spatio-temporal attention module for human action recognition. *IEEE Access* (99), 1 (2020)
13. Jegham, I., Khalifa, A.B., Alouani, I., et al.: Vision-based human action recognition: An overview and real world challenges. *Digit. Investig.* **32**(7), 200901 (2020)
14. Liu, S., Pan, Z., Cheng, X.: A novel fast fractal image compression method based on distance clustering in high dimensional sphere surface. *Fractals* **25**(4), 1740004 (2017)