

## Research Article

# Anomaly Recognition Algorithm for Human Multipose Motion Behavior Using Generative Adversarial Network

Nan Zhang,<sup>1</sup> Jie Ren,<sup>2</sup> Qixiao Xu ,<sup>1</sup> Hao Wu,<sup>1</sup> and Meng Wang<sup>1</sup>

<sup>1</sup>Physical Education Department, Beijing University of Technology, Beijing 100124, China

<sup>2</sup>College of Physical Education and Training, Harbin Sport University, Harbin, 150008 Heilongjiang, China

Correspondence should be addressed to Qixiao Xu; [xuqixiao@bjut.edu.cn](mailto:xuqixiao@bjut.edu.cn)

Received 22 November 2021; Accepted 3 January 2022; Published 17 January 2022

Academic Editor: Nima Jafari Navimipour

Copyright © 2022 Nan Zhang 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.

Human multipose motion behavior is similar; there are many actions. However, it is difficult to recognize abnormal behavior. The existing human motion behavior anomaly recognition methods have the problems of low accuracy and being time-consuming. Therefore, an anomaly recognition method of human multipose motion behavior using Generative Adversarial Network (GAN) is proposed. The Gauss model is used to segment the human multipose motion behavior image, and the image foreground of the segmentation result is the human motion target detection result. The Shi-Tomasi algorithm is selected to extract contour feature points from human motion object detection results. The extracted contour features are set as hidden space random variables and input into the GAN. The GAN uses the generator and discriminator to recognize the multipose human motion behavior and determine whether the multipose human motion behavior is abnormal. The results show that the proposed algorithm can accurately recognize abnormal human multipose motion behavior, the recognition accuracy is higher than 99%, and the average recognition time is less than 200 ms. The shadow removal effect of the foreground image obtained by the proposed algorithm can realize the accurate recognition of human multipose motion behavior abnormalities and provide a reliable basis for research in related fields.

## 1. Introduction

Human posture behavior recognition is a research in the field of video analysis. The goal of human posture behavior recognition is to classify the specific actions of humans [1]. Human abnormal behavior detection is a branch of human behavior recognition, that is, to recognize specific behaviors in different scenes, such as fighting in banks and other literature behaviors. Human posture behavior technology has been widely used in auxiliary medical treatment, video surveillance, and so on. The recognition of multipose motion behavior abnormalities needs to determine whether there are abnormalities by analyzing human posture [2].

The specific position and specific behavior of the motion target in a fixed time are analyzed through the abnormal recognition of multipose motion behavior [3, 4], so as to realize the automatic recognition and positioning of human posture. The abnormal recognition of human multipose motion behavior can be applied to many fields such as medical treat-

ment and security. Human motion posture has the features of high nonlinearity and degree of freedom [5, 6] and has high diversity and complexity. It is challenging to accurately recognize human multipose motion behavior abnormalities. The feature extraction of human motion posture is very important [7]. The accuracy of feature extraction determines the accuracy of multipose motion behavior. Feature extraction is realized by extracting texture features, color features, contour features, and other features of human motion posture [8, 9]. Human motion is formed by posture sequence, so it is more practical to extract human contour features from images.

The main contributions of this paper are as follows. (1) A new human multipose motion behavior anomaly recognition algorithm based on GAN is proposed to detect human motion targets in human multipose motion images. (2) Aiming at the problem of many similar actions of human multipose motion behavior, the human motion region and background region in the image are segmented, and the

human motion posture features are extracted from the human motion target region to reduce the difficulty of behavior anomaly recognition. (3) The GAN is used to obtain the abnormal recognition results of human multipose motion behavior. The network has the features of flexible application and strong learning ability. It is applied to the abnormal recognition of human multipose motion behavior to optimize the recognition performance.

## 2. Related Works

At present, there are many researches on motion recognition. Literature [10] applies convolutional neural network (CNN) to human motion recognition and proposes three different CNN structures. Firstly, four different information channels are generated from each frame in the horizontal and vertical directions through optical flow and gradient to be applied to three-dimensional (3D) CNN. Then, three architectures are proposed, namely, single stream, dual stream, and four stream 3D CNN. In the single stream model, four information channels per frame are applied to the single stream. In the dual flow structure, optical flow- $x$  and optical flow- $y$  are applied to one flow, and gradient- $x$  and gradient- $y$  are applied to another flow. In the four stream architecture, each information channel is applied to four independent streams, by evaluating the architecture of the action recognition system.

Literature [11] applied shape time dynamics to human motion recognition and proposed a human motion recognition framework with constant depth perspective, which is a new integration of two important motion clues: motion and shape time dynamics (STD). The motion flow encapsulates the motion content of the action into RGB dynamic image, which is generated by approximate rank pool (ARP) and processed by fine-tuning reception V3 model. The STD stream uses a series of long short-term memory (LSTM) and Bi LSTM learning models to learn the long-term view invariant shape dynamics of actions. Human pose model (HPM) generates view invariant features of human pose frames with key depth based on structural similarity index matrix (SSIM). According to three types of postfusion techniques, namely, maximum average and multiplication, the final prediction of single stream fraction is made. Literature [12] proposed an effective method for human motion recognition (HAR) from silhouette image sequences in video. The effectiveness of this method lies in feature extraction and action classification. The method includes scale translation normalization and distortion contour removal, which are used to extract the newly introduced spatiotemporal features, namely, active region energy feature (AREF), and trajectory analysis. In addition, the method uses low-dimensional eigenvectors, which makes the cost of the whole process effective in terms of time requirements. The results on the published Weizmann and Muhavi data sets clearly verify the efficiency of the proposed technology in the related work on the accuracy of human behavior detection. Literature [13] proposed a human behavior anomaly detection method based on the combination of deep learning and artificial features. Firstly, the key points of human 3D

skeleton in time series are extracted by Yolo V4, and the mean shift target tracking algorithm is applied. Then, the key points are transformed into spatial RGB and put into multilayer convolutional neural network for recognition, abnormal behaviors such as hitting, throwing, climbing, and approaching. Literature [14] uses deep learning technology, including CNN and LSTM network, to build a deep network in a multiperspective framework to learn the long-term correlation of human behavior recognition from video. Two cameras are used as sensors to effectively overcome the problems of occlusion and fuzzy contour and improve the accuracy and performance of multiview frame. After a series of image preprocessing on the original data, the human contour image is obtained as the input of the training model. Literature [15] combines spatiotemporal CNN with handmade feature sets for anomaly detection in continuous video frames. Handmade features learn sparse features extracted from moving human image units, including moving pixels to reduce computing costs. The CNN model architecture is used to extract spatiotemporal features and complete the recognition of human abnormal behavior. However, human multipose motion behavior is similar, and there are many actions. This leads to the question of low accuracy and being time-consuming in the application of the above existing methods.

## 3. Methodology

*3.1. Detecting Human Moving Targets.* The multipose motion image of human contains background factors such as shadow and illumination. Therefore, it is difficult to obtain human target from the image. The human motion target detection process for segmenting human motion target and image background is as follows.

*3.1.1. Establish Gaussian Model.* Let  $X_i$  be the pixel color of the random point in the human multipose motion image, and the probability density function expression of the pixel color is established as follows:

$$f(X_i = x) = \sum_{i=1}^k \varphi_{i,t} \cdot \alpha(x, \beta_{i,t}, \Sigma_{i,t}), \quad (1)$$

where  $\beta_{i,t}$  and  $\varphi_{i,t}$  are the mean and weight of Gaussian distribution when the time is  $t$ , respectively.  $\Sigma_{i,t}$  and  $\alpha(\cdot)$  are covariance matrix and Gaussian distribution, respectively. According to equation (1), the utilization number of pixel color  $X_i$  probability density function is  $k$ . When the time is  $t$ , the expression of the  $i$ th Gaussian distribution  $\alpha(\cdot)$  is as follows:

$$\alpha(x, \beta_{i,t}, \Sigma_{i,t}) = \frac{e^{-(1/2)(x_t - \beta_{i,t})^T \Sigma_{i,t}^{-1} (x_t - \beta_{i,t})}}{(2\pi)^{n/2} |\Sigma_{i,t}|^{1/2}}, \quad (2)$$

where  $n$  represents the dimension of pixel color  $X_i$ ,  $i = 1, 2, \dots, k$ .

**3.1.2. Update Model.** Let the random pixel value in the image be  $I_t$ . The expression that the Gaussian function matches the pixel value is as follows:

$$|I_t - \beta_{i,t-1}| \leq D \cdot \lambda_{i,t-1}. \quad (3)$$

The expression of each parameter update is as follows:

$$\begin{aligned} \lambda_{i,t}^2 &= (1 - \rho)\lambda_{i,t-1}^2 + \rho(I_t - \beta_{i,t-1})^2, \\ \beta_{i,t} &= (1 - \rho)\beta_{i,t-1} + \rho I_t, \\ \varphi_{i,t} &= (1 - a)\varphi_{i,t-1} + a, \end{aligned} \quad (4)$$

where  $a$  and  $\rho$  represent adjustable learning rate and parameter learning rate, respectively.

**3.1.3. Segmented Image.** Normalize all the weights obtained, and sort all Gaussian distributions  $\varphi_{i,t}/\lambda_{i,t}$  according to the order from large to small. The first  $M$  distributions to be sorted must meet the following conditions:

$$\sum_{k=i_1}^{i_M} \varphi_{i,t} \geq \varsigma. \quad (5)$$

The first  $M$  distributions are the background distribution of human multipose motion image,  $\varsigma$  is the weight threshold.

Randomly select pixels, which are foreground pixels when they can meet the established Gaussian mixture model [16]. When the pixel cannot meet the established Gaussian mixture model, this pixel is the background pixel [17]. Through the above process, the foreground and background of human multipose motion image are segmented, and the motion targets in human multipose motion image are detected.

**3.2. Extracting Contour Feature Points.** After segmenting the foreground and background of human multipose motion image, it is necessary to extract the contour feature points in the foreground of human motion image, so as to provide the basis for human multipose motion behavior anomaly recognition based on GAN. Harris corner detection is the most widely used image contour feature extraction method. A corner is the intersection of two edges. However, Harris corner detection needs to calculate the empirical constant  $k$ , and the operation process has high complexity. Therefore, the Shi-Tomasi algorithm is proposed. Since the stability of the Harris corner detection algorithm is related to the value of the empirical constant  $k$  and the empirical constant  $k$  is an empirical value, it is difficult to set the optimal value. Shi-Tomasi found that the stability of corner points is actually related to the smaller eigenvalue of the matrix, so the smaller eigenvalue is directly used as the score, so there is no need to adjust the  $k$  value. Therefore, the application of Shi-Tomasi is simpler, and the image contour feature points obtained by Shi-Tomasi are also very accurate.

The Shi-Tomasi algorithm is selected to extract contour feature points in human motion image.  $(x, y)$  and  $I(x, y)$  are the random point coordinates of the multipose motion

grayscale image of the human and the gray value of the point, respectively. The point  $(x, y)$  as the center point is using window  $S$  is established. The window size is  $n \times n$ , window  $[x, y]$  is translated, and the gray value changes as follows after translation:

$$L(\Delta x, \Delta y) = \sum_{(x,y) \in S} \phi(x, y) [I(x + \Delta x, y + \Delta y) - I(x, y)]^2. \quad (6)$$

Taylor expansion process  $I(x + \Delta x, y + \Delta y)$ , delete the terms higher than the second order, and the expression is as follows:

$$L(\Delta x, \Delta y) = \Delta x^2 \sum_{(x,y) \in S} \phi(x, y) I_x^2 + 2\Delta x \Delta y \sum_{(x,y) \in S} I_x I_y + \Delta y^2 \sum_{(x,y) \in S} \phi(x, y) I_y^2, \quad (7)$$

where  $I_x$  and  $I_y$  are the partial derivatives of human multipose motion images and  $\phi(x, y)$  is the Gaussian filter.

Equation (7) can be transformed into the following matrix:

$$L(\Delta x, \Delta y) = [\Delta x, \Delta y] M \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}. \quad (8)$$

The expression in equation (8) is as follows:

$$M = \sum_{(x,y) \in S} \phi(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}. \quad (9)$$

The Shi-Tomasi algorithm extracts contour feature points from human multipose motion images. The extraction rules are as follows:

- (1) Randomly select the pixel point, let the pixel point be the midpoint, establish a window with size  $n \times n$ , and move the window in different directions [18]
- (2) Compare the changes of gray value in human multipose motion images. When the window is moved in different directions, when the gray value in the image is fixed, it means that the area is flat [19], and there is no feature point; when moving along a fixed direction and there is only a small change in the gray value, it means that the area is a straight line area
- (3) When the established window moves in a random direction [20], when the gray level in the multipose motion image of the human changes greatly, the center point of the established window is the feature point of the human motion image, and this feature point can be used to realize the anomaly recognition of human multipose motion behavior

**3.3. Anomaly Recognition of GAN.** The discriminator in the GAN is used for the final recognition of human multipose motion behavior. The discriminator has the function of classification [21] and can judge whether the input sample data

belongs to the generated sample data or real data. GAN is an important GAN algorithm with high linear correlation performance. When the GAN recognizes the multipose movement behavior of human, it can quickly obtain the recognition results [22]. There is no need to interfere with the learning process of the GAN, and accurate recognition results can be obtained only by evaluating the model effect after learning. The structure of GAN is shown in Figure 1.

$x$  and  $z$  represent the sample data and random noise of the input GAN, respectively, and  $A$  and  $B$  represent the generator and discriminator of the GAN, respectively. The expression of the GAN is as follows:

$$\min_A \max_B V(B, A) = E_{x \sim P(x)} [B(x)] + E_{z \sim P(z)} [\lg(1 - B(A(z)))], \quad (10)$$

where  $\lg(1 - B(A(z)))$  and  $\lg B(x)$  are generated data and discriminator judgment, respectively, and  $V$  and  $E$  represent objective function and discriminator function, respectively.

The discriminator and the generated data play games with each other by using the maximum and minimum values, and the optimization of parameter  $B$  and parameter  $A$  is completed through multiple iterations until the discriminator and the generated data are in Nash equilibrium.

The GAN is prone to gradient dispersion in the established minimization objective function. Therefore, the objective function cannot update the generator in the GAN [23], which will reduce the stability of the GAN. LSGAN method is selected to solve the training instability of GAN. LSGANs ensure the classification accuracy of GAN through cross-entropy. The objective function expression of GAN discriminator optimized by LSGANs is as follows:

$$\min_B E_{x \sim P_{\text{data}}(x)} [(B(x) - b)^2] + E_{z \sim P_Z(z)} [(B(A(z)) - a)^2]. \quad (11)$$

The objective function expression of GAN generator is as follows:

$$E_{z \sim P_Z(z)} [(B(A(z)) - c)^2]. \quad (12)$$

After completing the antagonistic generation network training, the adjusting parameters  $a$ ,  $b$ , and  $c$  need to meet the following conditions:

$$\begin{cases} b - c = 1, \\ b - a = 2. \end{cases} \quad (13)$$

Through the above process, the uncertainty caused by the training process of GAN can be improved [24], the diversity of GAN is effectively improved, and the recognition accuracy of human multipose motion behavior abnormalities is improved.

The GAN is applied to the abnormal recognition of human multipose motion behavior. The process of abnormal recognition is shown in Figure 2.

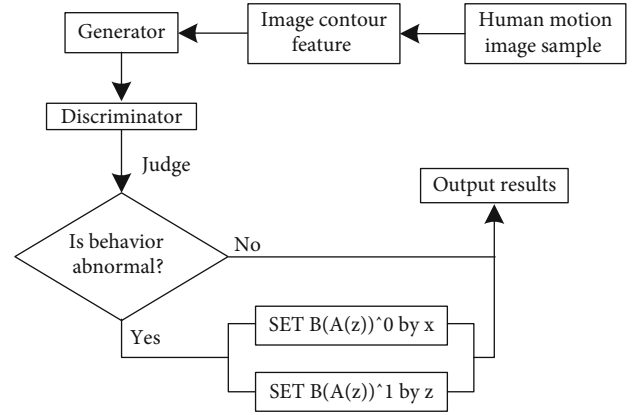


FIGURE 1: Structure of abnormal recognition behavior of GAN.

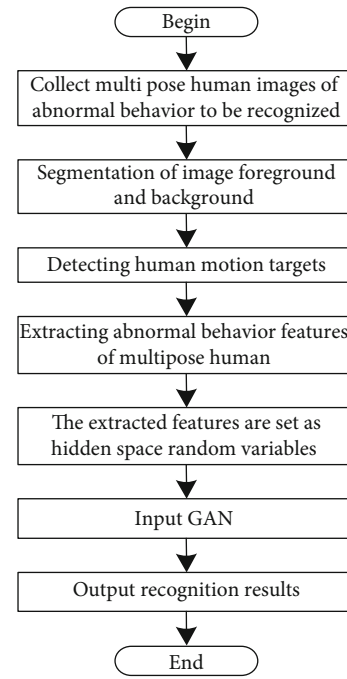


FIGURE 2: Abnormal recognition process of human multipose motion behavior.

According to Figure 2, firstly, the multipose human image with abnormal behavior to be recognized is segmented. After the image segmentation is completed, the detection results of human motion targets are obtained, the contour features of human motion behavior are extracted from the extracted human motion targets, and the extracted features are set as hidden space random variables and input into the GAN. The GAN outputs the abnormal behavior recognition results of multipose human motion behavior to determine whether the multipose human motion behavior is abnormal.

**3.4. Data Sets and Evaluating Index.** In order to verify the effectiveness of the proposed algorithm in recognizing abnormal human multipose motion behavior by GAN, the Occlusion\_person 3D data set and CMU panoptic data set are selected as the test data set. Occlusion\_person 3D data

set has 4.8 million 3D human postures and corresponding images, a total of 200 experimenters, and a total of 23 action scenes. CMU panoptic data set is produced by CMU University and collected by 480 VGA cameras, 30+ HD cameras, and 10 kinect sensors. The above two data sets are typical human posture data sets. The two data sets include walking, running, kicking, jumping, standing, squatting, hands up, reverse, head down, head up, and other postures, which are extremely typical. Forty thousand images were randomly selected from the two data sets. In this experiment, twenty thousand images are selected for data training, and the remaining half are used for experimental test and analysis. The number of images of each behavior posture is shown in Table 1. In the simulation platform, the abnormal behavior of the image is recognized.

The accuracy, precision, and recall are selected as the evaluation indexes to evaluate recognition performance of the proposed algorithm. The calculation expressions of three indexes are as follows:

$$\begin{aligned} \text{accuracy} &= \frac{r_p + r_n}{r_p + r_n + h_p + h_n}, \\ \text{precision} &= \frac{r_p}{r_p + h_p}, \\ \text{recall} &= \frac{r_p}{r_p + h_n}, \end{aligned} \quad (14)$$

where  $r_p$  and  $r_n$  represent the number of abnormal behaviors recognized by the algorithm as abnormal behaviors and nonabnormal behaviors, respectively.  $h_p$  and  $h_n$  are the numbers of abnormal motion behaviors and nonabnormal behaviors, respectively. The accuracy and precision are used to measure the recognition level of the algorithm and the level that can avoid misrecognition, and the recall rate reflects the level at which the algorithm can correctly classify whether the behavior is abnormal sports behavior.

#### 4. Results and Discussion

The proposed algorithm is used to recognize multipose human motion behavior, and the abnormal behavior recognition results of the proposed algorithm are compared with actual abnormal behavior results, as shown in Table 2.

According to Table 2, for the 10 motion behaviors, the total number of images is between 168 and 667, and the number of abnormal images is between 19 and 63. The proposed method is used to recognize these images, and the results are close to the actual number of abnormal images. The recognition rate of walking, jumping, standing, squatting, and hands up can be 100%. The data shows that the proposed algorithm used to recognize abnormal human multipose motion behavior and the actual abnormal human multipose motion behavior is very small; it shows that the proposed algorithm has high effectiveness in recognizing abnormal human multipose motion behavior.

TABLE 1: Image sample set attributes.

Type	Resolution (dpi)	Space size (kb)
Walking	120	11779
Running	140	7354
Kicking	80	20374
Jumping	70	5216
Standing	100	12538
Squatting	68	4297
Hands up	100	7541
Reverse	50	1569
Head down	80	20685
Head up	67	6328

TABLE 2: Abnormal recognition results.

Motion behavior	Total number of images (piece)	Actual abnormal quantity (piece)	Abnormal recognition quantity (piece)
Walking	351	38	38
Running	254	48	47
Kicking	184	52	51
Jumping	645	34	34
Standing	284	19	19
Squatting	168	27	27
Hands up	294	34	34
Reverse	587	29	28
Head down	667	47	46
Head up	566	63	62

The proposed algorithm randomly divides one of the images and detects the result of human motion targets, as shown in Figure 3.

According to Figure 3, the proposed algorithm can achieve accurate detection of human motion targets in multipose human motion images. The proposed algorithm has a high level of image segmentation. It can accurately extract human motion targets through effective image segmentation ability, which provides the basis for accurate recognition of abnormal motion behavior.

The proposed algorithm is compared with the algorithm in literature [10–14]. The comparison results of the accuracy and recall of the six algorithms for the abnormal recognition of multipose motion behavior in the experimental data set are shown in Figures 4 and 5. The squatting posture is selected for the same posture test.

According to Figures 4 and 5, under different postures, the accuracy of the proposed algorithm is always higher than 99%. For the abnormal behavior recognition of the same posture, the accuracy always fluctuates between 99% and 100%. In contrast, in the process of different posture recognition, the recognition accuracy of literature [10] algorithm, literature [11] algorithm, literature [12] algorithm, literature



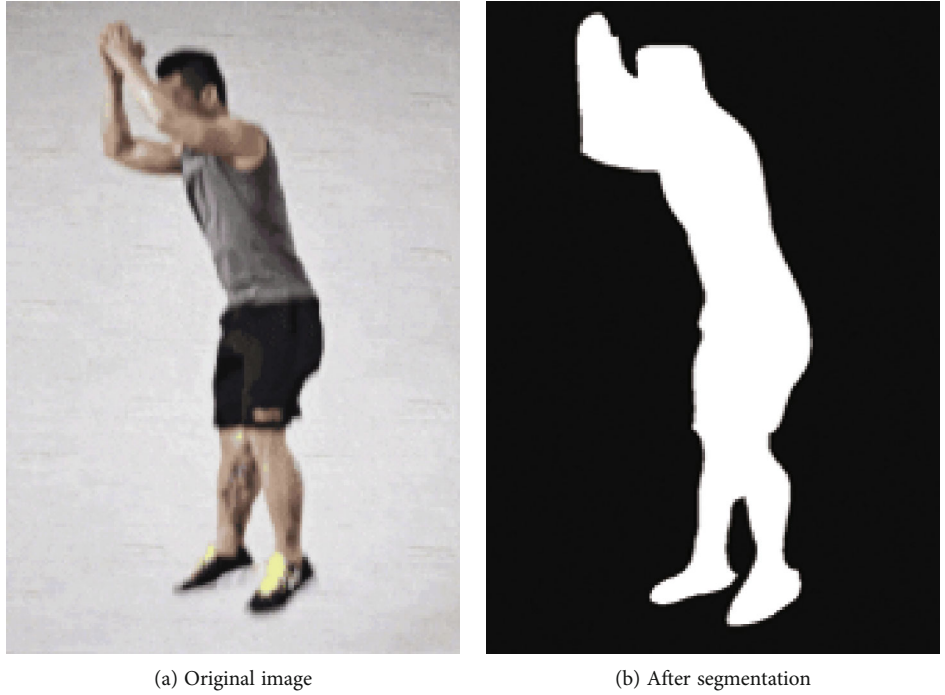


FIGURE 3: Human motion target extraction results.

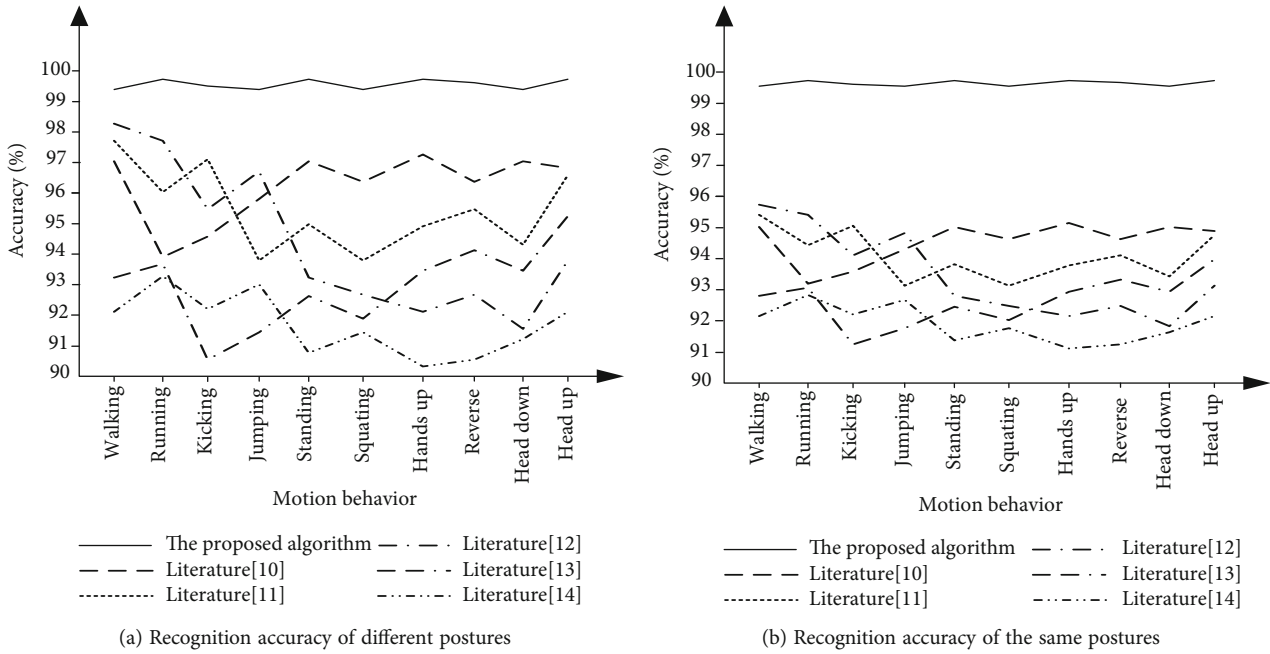


FIGURE 4: Comparison results of recognition accuracy.

[13] algorithm, and literature [14] algorithm fluctuates between 90% and 98.5%, and the recognition accuracy of the same posture is always lower than 96%. The recognition accuracy of the proposed algorithm is significantly higher than that of the other five algorithms. In addition, the recall rate of the proposed algorithm for recognizing multipose and the same posture human motion behavior abnormality

is higher than 99%, and the recall rate of literature algorithm for recognizing multipose human motion behavior abnormality is lower than the proposed algorithm. The comparison results verify that the proposed algorithm has high performance of human motion behavior anomaly recognition. The proposed algorithm effectively segments the multipose human motion behavior image, selects the GAN, and

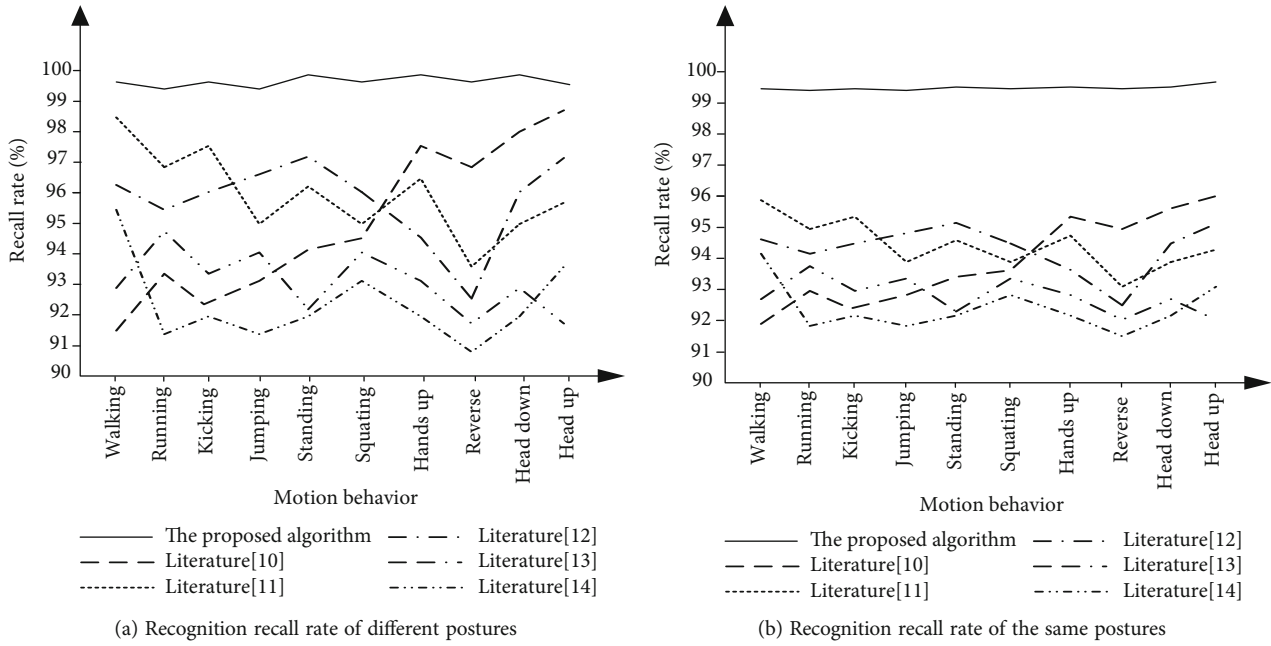


FIGURE 5: Comparison results of recognition recall rate.

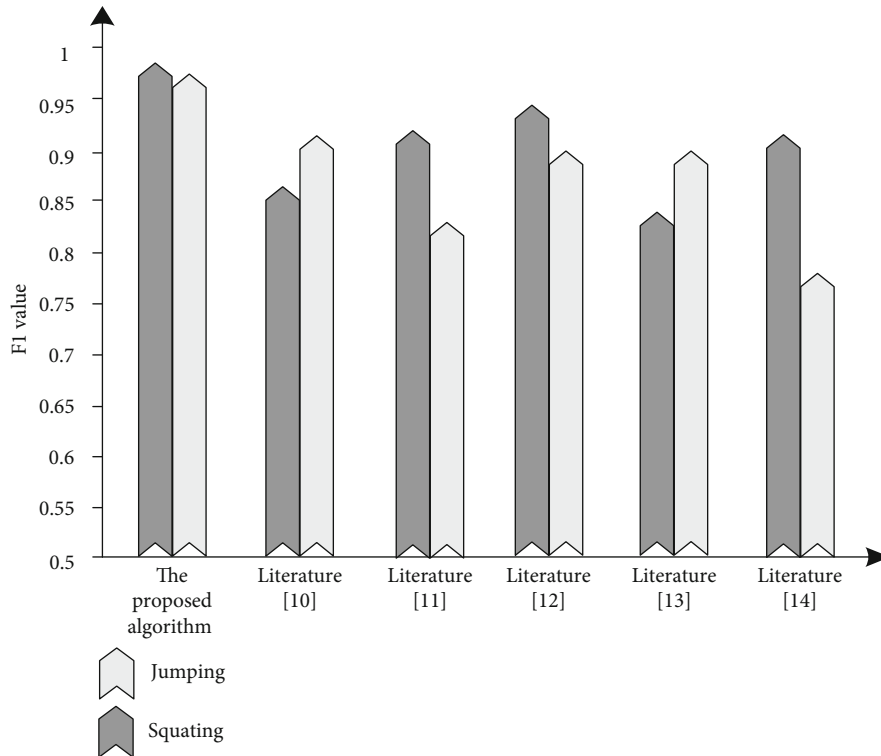


FIGURE 6: Comparison results of  $F1$  value.

uses the segmented image to realize the accurate recognition of multipose human motion behavior anomaly.

In order to further measure the recognition performance of this algorithm,  $F1$  value is selected as the test index to measure the recognition of abnormal human multipose motion behavior. Six algorithms are used to recognize

the  $F1$  value of abnormal human multipose motion behavior in two data sets. The  $F1$  value is affected by the accuracy rate and recall rate. The higher the accuracy rate is, the higher the recall rate is, and the better the  $F1$  index is. The statistical results of  $F1$  value indexes are shown in Figure 6.

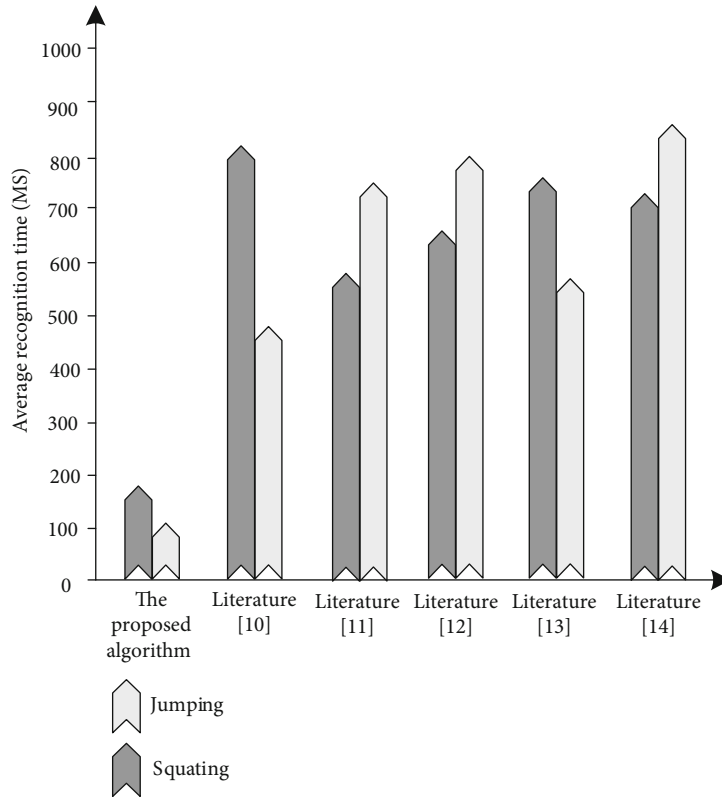


FIGURE 7: Comparison results of recognition time.

As can be seen in Figure 6, the  $F1$  value of the proposed algorithm to recognize the abnormal multipose motion behavior of humans is higher than that of the other five algorithms. It is verified again that the proposed algorithm has high recognition performance. This is because this algorithm uses Gaussian model to segment human multipose motion behavior image, and the image foreground of the segmentation result effectively reduces the difficulty of image segmentation, improves the recognition accuracy of human posture behavior, and makes  $F1$  value higher.

The real-time performance of the recognition algorithm is very important. Six algorithms are used to recognize the recognition time of abnormal human multipose motion behavior. The statistical results are shown in Figure 7.

As can be seen in Figure 7, the average recognition time of using the proposed algorithm to recognize human multipose motion behavior abnormalities is less than 200 ms, which verifies that the proposed algorithm has high recognition real-time performance. This is because the proposed algorithm selects the Shi-Tomasi algorithm to extract the contour feature points in the human motion image from the human motion target detection results. Based on this, the extracted contour features are set as hidden space random variables and input into the GAN to determine whether the multipose human motion behavior is abnormal. GAN is a very mature technology for enhancing image quality and has application advantages. The proposed algorithm has a wide range of applications, and the real-time performance of the algorithm is very important. The proposed algorithm

has high real-time recognition performance, which can improve the application performance of the algorithm.

## 5. Conclusions

GAN is applied to human multipose motion behavior anomaly recognition. Firstly, the multipose human motion image is segmented to obtain the human motion target area in the image, and the human motion contour features are extracted from the human motion target area. Based on the GAN, the extracted features are used to realize the abnormal recognition of human multipose motion behavior. Experiments show that the recognition results of human multipose motion behavior abnormalities have high accuracy, and the proposed algorithm has high feasibility. The proposed algorithm can not only realize the abnormal recognition of human multipose motion behavior and ensure the recognition accuracy but also has the advantage of high real-time recognition. The proposed algorithm solves the defect of low recognition performance caused by too few sample data in the past.

However, for specific occasions, the multipose behavior of humans is greatly affected by light, resulting in great difficulty in image acquisition, and the collected image is prone to distortion. In the practical application of human behavior recognition method, the combination of light compensation technology and motion target detection algorithm to enhance the effectiveness of data set still needs further research.



## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there is no conflict of interest with any financial organizations regarding the material reported in this manuscript.

## References

- [1] J. S. Park, C. Park, and D. Manocha, "I-Planner: intention-aware motion planning using learning-based human motion prediction," *The International Journal of Robotics Research*, vol. 38, no. 1, pp. 23–39, 2019.
- [2] W. Wang, F. Bu, Z. Lin, and S. Zhai, "Learning methods of convolutional neural network combined with image feature extraction in brain tumor detection," *Access*, vol. 8, pp. 152659–152668, 2020.
- [3] K. Yildiz, "Identification of wool and mohair fibres with texture feature extraction and deep learning," *IET Image Processing*, vol. 14, no. 2, pp. 348–353, 2020.
- [4] S. Qiao, L. Wang, and Z. Gao, "Group behavior recognition based on deep hierarchical network," *Neural Computing and Applications*, vol. 32, no. 10, pp. 5389–5398, 2020.
- [5] N. Wei, "Research on the algorithm of painting image style feature extraction based on intelligent vision," *Future Generation Computer Systems*, vol. 123, no. 24, pp. 196–200, 2021.
- [6] M. Tan, J. Zhou, Z. Peng, Y. Jun, and F. Tang, "Fine-grained image classification with factorized deep user click feature," *Information Processing and Management*, vol. 57, no. 3, 2020.
- [7] A. Safonova, E. Guirado, Y. Maglinets, D. Alcaraz-Segura, and S. Tabik, "Olive tree biovolume from UAV multi-resolution image segmentation with Mask R-CNN," *Sensors*, vol. 21, no. 5, p. 1617, 2021.
- [8] S. Prasath, R. Pelapur, G. Seetharaman et al., "Multiscale structure tensor for improved feature extraction and image regularization," *IEEE Transactions on Image Processing*, vol. 28, no. 12, pp. 6198–6210, 2019.
- [9] Y. Zhou and Z. Gao, "Intelligent recognition of medical motion image combining convolutional neural network with Internet of things," *IEEE Access*, vol. 7, pp. 145462–145476, 2019.
- [10] H. V. A. Chenarlog and F. Razzazi, "Multi-stream 3D CNN structure for human action recognition trained by limited data," *IET Computer Vision*, vol. 13, no. 3, pp. 338–344, 2019.
- [11] C. Dhiman and D. K. Vishwakarma, "View-invariant deep architecture for human action recognition using two-stream motion and shape temporal dynamics," *IEEE Transactions on Image Processing*, vol. 29, pp. 3835–3844, 2020.
- [12] S. Maity, A. Chakrabarti, and D. Bhattacharjee, "Robust human action recognition using AREI features and trajectory analysis from silhouette image sequence," *IETE Journal of Research*, vol. 65, no. 2, pp. 236–249, 2019.
- [13] S. Xie, X. Zhang, and J. Cai, "Video crowd detection and abnormal behavior model detection based on machine learning method," *Neural Computing and Applications*, vol. 31, no. S1, pp. 175–184, 2019.
- [14] Y. L. Hsueh, W. N. Lie, and G. Y. Guo, "Human behavior recognition from multiview videos," *Information Sciences*, vol. 517, pp. 275–296, 2020.
- [15] B. Sabzalian, H. Marvi, and A. Ahmadyfard, "Deep and sparse features for anomaly detection and localization in video," in *2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA)*, pp. 173–178, Tehran, Iran, 2019.
- [16] A. Soualmi, A. Alti, and L. Laouamer, "A novel blind medical image watermarking scheme based on Schur triangulation and chaotic sequence," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 1, 2022.
- [17] E. Cuevas, H. Becerra, A. Luque, and M. A. Elaziz, "Fast multi-feature image segmentation," *Applied Mathematical Modelling*, vol. 90, no. 5, pp. 742–757, 2021.
- [18] T. Santos, S. Schrunner, B. C. Geiger et al., "Feature extraction from analog wafermaps: a comparison of classical image processing and a deep generative model," *IEEE Transactions on Semiconductor Manufacturing*, vol. 32, no. 2, pp. 190–198, 2019.
- [19] P. Zhang, H. He, and L. Gao, "A nonlinear and explicit framework of supervised manifold-feature extraction for hyperspectral image classification," *Neurocomputing*, vol. 337, pp. 315–324, 2019.
- [20] Z. Xia, J. Xing, C. Wang, and X. Li, "Gesture recognition algorithm of human motion target based on deep neural network," *Mobile Information Systems*, vol. 2021, Article ID 2621691, 12 pages, 2021.
- [21] N. Martins, "Adversarial machine learning applied to intrusion and malware scenarios: a systematic review," *Access*, vol. 8, pp. 35403–35419, 2020.
- [22] P. Dasgupta and J. B. Collins, "A survey of game theoretic approaches for adversarial machine learning in cybersecurity tasks," *AI Magazine*, vol. 40, no. 2, pp. 31–43, 2019.
- [23] S. G. Finlayson, J. D. Bowers, J. Ito, J. L. Zittrain, A. L. Beam, and I. S. Kohane, "Adversarial attacks on medical machine learning," *Science*, vol. 363, no. 6433, pp. 1287–1289, 2019.
- [24] D. Roy, T. Mukherjee, and M. Chatterjee, "Machine learning in adversarial RF environments," *IEEE Communications Magazine*, vol. 57, no. 5, pp. 82–87, 2019.