

Abnormal Behavior Detection in Surveillance Video via Multi-Input Feature Clustering with GAN-Augmented Autoencoders

Huiying Li¹, Liping Wang^{2*}, Yongna Jiao³

¹Department of Humanities and Social Sciences, Shijiazhuang Institute of Railway Technology, Shijiazhuang, 050062, China

²Department of Tourism and Business, Handan Polytechnic College, Handan, 056004, China

³Office of Students, Shijiazhuang Institute of Railway Technology, Shijiazhuang, 050062, China

E-mail: 13831028868@163.com

Keywords: abnormal behavior detection, monitoring videos, tourism scenic area, clustering algorithm, mixed multi-input features, generative adversarial network

Received: May 22, 2025

With the booming development of the global tourism industry, the increase in tourists has gradually made the safety management of tourist attractions more important. Monitoring abnormal behavior in tourist attractions is crucial in the safety management. To improve the accuracy of monitoring abnormal behavior in tourist attractions, this study combines convolutional neural networks with autoencoder network structures to reduce the learning generalization ability of convolutional neural networks. Attention mechanism is incorporated to improve sensitivity and recognition accuracy of abnormal behavior in complex environments. The method was experimentally validated using the CUHK Avenue and UCSD datasets, and compared with existing baseline methods. The results showed that the mixed multi-input feature clustering algorithm based on deep convolutional autoencoder had better detection performance than traditional methods on these two datasets. On the CUHK Avenue dataset, the AUC value was 91.9%, which was 27.1%, 10.6%, 15.0%, and 2.8% higher than that of the Adam, MDT, SF, and SRC methods, respectively. On the UCSD dataset, the AUC value reached 94.7%, which was 31.0% higher than that of the other four methods. In addition, the precision on the CUHK Avenue dataset was 94.5%, the recall rate was 95.6%, and the error rate was 12.6%. On the UCSD dataset, the precision was 95.2%, the recall rate was 94.8%, and the error rate was 10.9%. Overall, the research on the detection method of abnormal behavior in tourist attraction monitoring videos based on mixed multi-input feature clustering algorithm has high detection accuracy and can provide more effective technical support for the safety management of tourist attractions.

Povzetek: DCAMMFCA združi SSD, pozornostno izboljšan konvolucijski avtoenkoder z GAN ter K-means gručenje mešanih časovno-prostorskih značilk za odkrivanje anomalij v turističnem nadzoru.

1 Introduction

As the economy and culture rapidly develop and the global tourism industry prospers, tourism has become an important venue for economic and cultural exchanges. Meanwhile, as modern cities continue to advance, the requirements for safety supervision in the public sector are also increasing. The monitoring system, as a key technology for security monitoring, has seen an increasing demand for its intelligence and information security [1-2]. Tourism Scenic Area (TSA) often faces challenges such as high pedestrian traffic and complex terrain, and traditional manual monitoring technologies often encounter high false positive rates [3]. Traditional video surveillance mainly relies on simple motion detection or algorithms with specific rules, such as fixed area intrusion detection and trajectory anomaly analysis. These monitoring technologies are effective enough in simple environments, but their effectiveness is limited when faced with dynamic and complex tourism scenes

[4]. For example, factors such as fluctuations in crowd density, environmental obstructions, and changes in lighting conditions can affect the accuracy of video detection [5]. In addition, due to the lack of intelligent factors, traditional video surveillance technology cannot effectively classify and store recorded data, resulting in huge data processing time and difficulty in obtaining all information. Therefore, an innovative approach based on the Mixed Multi-input Feature Clustering Algorithm (MMFCA) is proposed for abnormal behavior detection on surveillance videos to address the low detection accuracy in video frame prediction and reconstruction in complex environments. Meanwhile, the optimized autoencoder based on attention mechanism is used as a Generative Adversarial Network (GAN) for feature extraction to improve sensitivity and recognition accuracy for abnormal behavior in complex environments.

The core question of the research is: "Can the combination of SSD-based spatial feature extraction and Time GAN attention autoencoder improve the accuracy

of abnormal behavior detection in TSA scenarios?” To verify this hypothesis, corresponding experiments are designed and various baseline methods are compared. The research hypothesis suggests that the Mixed Multi-input Feature Clustering Algorithm based on Deep Convolutional Autoencoder (DCAMMFCA) can effectively improve the abnormal behavior detection in scenic surveillance videos, especially when dealing with small object detection in complex environments. The research is divided into six sections. The first section is the introduction. The second section reviews the current research status of intelligent monitoring systems and abnormal behavior detection both domestically and internationally. Next, the third section introduces a monitoring video anomaly detection method based on the DCAMMFCA. The fourth section analyzes the abnormal behavior detection results based on this algorithm and compares them with existing methods. The fifth section is discussion. The sixth section is the conclusion.

2 Related works

As an important research direction in computer vision, intelligent monitoring systems have received attention from many experts and scholars and have achieved many results. Jenssen et al. proposed an automatic vision-based power line inspection and monitoring system to monitor power lines. This system utilized deep learning technology for network construction and utilized deep residual network structure for damage monitoring of power line components. These results confirmed that the method had high monitoring accuracy [6]. Yousefi et al. proposed a monitoring system that combined sensor systems for real-time monitoring of food in the production chain. This design utilized biosensors for monitoring production environment humidity, temperature, and gases. These results confirmed that this method monitored food quality and ensured food production safety [7]. Pimenov et al. combined artificial intelligence technology with sensors to design a monitoring system for real-time monitoring during tool processing. This system could monitor the real-time status of cutting tools during machining operations and utilize machining responses to monitor the surface roughness of the tools. These results confirmed that this method effectively improved dimensional accuracy and production efficiency during the machining process [8]. Liu combined machine learning technology with data mining technology for real-time monitoring of abnormal advertisements to maintain the integrity and efficiency of advertising campaigns. The results showed that this method could monitor various measures of advertising activities in a vigilant manner and was feasible [9]. Mattera et al. developed a line arc additive manufacturing program using artificial intelligence technology to monitor the production process of arc

additive manufacturing. The program included a defect detection module that could monitor the production process of arc additive manufacturing. The results showed that this method was helpful for parameter control in the manufacturing process [10].

Abnormal behavior detection plays an important role in intelligent monitoring. ALDHAMARI et al. put forward a high-performance structure to design a smart monitoring system with human behavior detection and classification. This framework utilized foreground optical flow energy to extract descriptive spatiotemporal features from surveillance videos. The orthogonal matching tracking algorithm was used to recover high-dimensional sparse features. These results confirmed that the method effectively improved the behavior detecting and classifying accuracy [11]. Hu et al. proposed a deep learning-based driver abnormal behavior detection system to effectively identify abnormal driver behavior. The system utilized stacked sparse autoencoders to learn driving behavior features, and then used greedy layering for training. These results confirmed that the method had high detection accuracy in detecting abnormal driving behavior [12]. Feizi et al. proposed a new normal behavior estimation model to accurately define abnormal behavior. This design utilized the histogram of directional optical flow as the basic local feature and utilized spectral clustering for similar feature clustering. These results confirmed that this method could effectively distinguish different behaviors [13]. Zhang et al. proposed a cloud platform virtual machine abnormal behavior monitoring system to improve the security and reliability of virtual machines. This system utilized incremental clustering algorithm for load information monitoring and local outlier factor algorithm for online anomaly detection. These results confirmed that this method could meet the real-time monitoring requirements [14]. Gao et al. used wireless sensors and discrete-time Markov chains to construct a user activity monitoring model connected to the medical Internet of Things for detecting abnormal behavior in patients with Alzheimer's disease. This model classified users' daily behaviors using probability calculation tree logic. The results showed that this method was feasible [15]. To monitor Ethereum fraud, Tan et al. proposed a method for mining Ethereum transaction records to monitor fraudulent transactions. This method used web crawling technology to obtain Ethereum addresses with fraud tags, and then used network embedding algorithms to extract node features for subsequent fraud transaction recognition. Finally, a graph Convolutional Neural Network (CNN) was used to classify the identified addresses. The results showed that the accuracy of Ethereum fraud transaction monitoring was as high as 96% [16]. The summary of relevant work is shown in Table 1.

Table 1: Summary of related work

Method	Dataset	Feature extraction method	Accuracy	Advantages	Limitations
Jenssen et al.	Power line dataset	Deep residual network	High accuracy	High monitoring precision	Only applicable to power line monitoring, not suitable for general scenarios
Yousefi et al.	Food production dataset	Biosensors	No clear accuracy data	Real-time monitoring of environmental data	Limited by environmental sensors, cannot handle dynamic behavioral changes
Pimenov et al.	Manufacturing dataset	Sensors + AI	No clear accuracy data	Improves production efficiency and accuracy	Specific to certain tools and processes, cannot generalize to other fields
Liu et al.	Advertising dataset	Machine learning + data mining	No clear accuracy data	Real-time monitoring of advertising activities	Cannot handle large-scale advertising data and rapidly changing behavioral patterns
Mattera et al.	Arc additive manufacturing dataset	AI + Sensors	No clear accuracy data	Good monitoring capabilities for manufacturing processes	Focused on additive manufacturing, cannot be generalized to other industries
ALDHA MARI et al.	Surveillance video dataset	Optical flow feature extraction + Orthogonal Matching Pursuit Algorithm	No clear accuracy data	Improves behavior classification accuracy	Only suitable for video surveillance, not applicable to other types of data
Hu et al.	Driver dataset	Sparse autoencoder	High accuracy	Detects driver abnormal behavior	Limited to driving behavior, not adaptable to other types of anomaly detection
Feizi et al.	Unknown dataset	Directional optical flow + spectral clustering	No clear accuracy data	Effectively distinguishes different behaviors	Possibly limited by specific behavior estimations
Zhang et al.	Virtual machine dataset	Incremental clustering + local outlier factor	High accuracy	Meets real-time monitoring requirements	Specific to virtual machine data, cannot handle other types of data
Gao et al.	Alzheimer's patients dataset	Wireless sensors + Markov chain	No clear accuracy data	Real-time monitoring of patient behavior	Only applicable to specific patient groups, cannot generalize
Tan et al.	Ethereum transaction records	Network embedding algorithm + Graph Convolutional Network	96%	High accuracy	Only applicable to Ethereum fraud monitoring

As shown in Table 1, these methods have failed to achieve their goals in the TSA context. For example, the power line monitoring method proposed by Jenssen et al. only focuses on a single domain and cannot cope with the changing monitoring scenarios. The proposed solution in this study has strong adaptability and can handle video surveillance in various environments. In addition, the food production chain monitoring method proposed by Yousefi et al. does not consider behavioral patterns and dynamic detection. The solution proposed in

this article, combined with deep learning technology, can dynamically identify and analyze abnormal behaviors.

In summary, many achievements have been made in research related to intelligent monitoring systems and abnormal behavior detection. However, there is still relatively little research on using feature storage autoencoders as network architectures for feature extraction and integrating multiple input features for clustering analysis to detect abnormal behavior in videos. The feature storage module, as an innovative method, is

introduced into the generator of the GAN for more efficient extraction and storage of multidimensional features. Therefore, the feature storage module optimizes the feature extraction process of the generator in abnormal behavior detection by storing and matching feature vectors to effectively improve the accuracy and sensitivity of detection.

3 Detection of abnormal behavior in monitoring videos based on deep convolutional autoencoder mixed multi-input feature clustering algorithm

A deep convolutional autoencoder network is used for abnormal behavior detection, and a clustering algorithm with mixed multi-input features is combined to improve

the detection performance of small targets. To reduce the impact of complex environments on feature extraction, the study further utilizes SSD object detection models to extract foreground targets in monitoring video images.

3.1 Abnormal behavior detection algorithm based on deep convolutional autoencoder network structure

CNN is a powerful tool specifically designed for analyzing visual images in deep learning. Through multi-level structural design, it can automatically and effectively learn spatial level features from image data [17-18]. In computer monitoring video analysis, since video images are essentially digital information having many pixels, it is particularly crucial to utilize CNN to extract the features of these images. The overall architecture of the research method is shown in Figure 1.

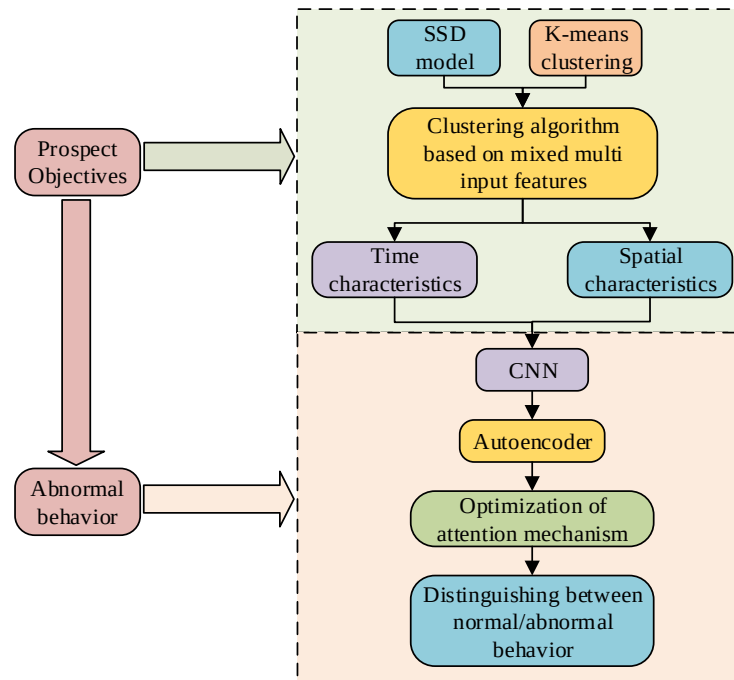


Figure 1: Overall architecture diagram of the research method

As shown in Figure 1, the research method combines the SSD object detection model with the K-means clustering algorithm to construct a MMFCA for foreground targets in complex environments. This algorithm can extract foreground targets from surveillance video images, classify the extracted features using the K-means clustering algorithm, and finally obtain the temporal and spatial features of the surveillance image. Next, the time and spatial features extracted from the foreground target are input into the CNN for abnormal behavior recognition. To address the strong generalization ability of CNN, this study combines CNN with autoencoder models and optimizes the model using attention mechanisms to improve the accuracy of distinguishing normal and abnormal behavior. There are three main basic structures of CNN. The input layer processes the original pixel data and converts it into

a form that the network can process. The feature extraction layer is usually composed of multiple alternating convolutional and pooling layers. The convolutional layer is responsible for extracting local features from the image. The pooling layer is responsible for down-sampling, reducing computational complexity while maintaining spatial hierarchy of features [19-20]. Finally, the fully connected layer maps the learned features to the final output. The function of the convolutional layer is represented by equation (1).

$$x_j^l = f \left(\sum_{i=1}^{N^{l-1}} G_{i,j}^l (k_{i,j}^l \otimes x_i^{l-1}) + b_j^l \right) \quad (1)$$

In equation (1), x_j^l represents the input function of the l -th convolutional layer, $\otimes$ means the convolution operation. b refers to the bias parameter. k is the

convolution kernel [21]. The mathematical expression for the pooling layer is represented by equation (2).

$$x_j^l = p(x_j^{l-1}) \quad (2)$$

In equation (2), $p(x)$ refers to pooling operation. The mathematical expression for the fully connected layer is represented by equation (3).

$$x^l = f(\omega^l x^{l-1} + b^l) \quad (3)$$

In equation (3), $f(x)$ refers to the nonlinear activation function. ω means the weight. Common nonlinear activation functions include Tanh, Sigmoid, and ReLu. The study uses the Sigmoid function as the activating function, represented by equation (4) [22].

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

Autoencoder is an unsupervised feature learning algorithm implemented through neural networks, whose core function is data dimensionality reduction and feature extraction [23]. Autoencoder learns data by encoding input data into a low-dimensional space, which is then decoded back to the original data. In this process, the autoencoder is to minimize the difference between input and output, which is also known as reconstruction error [24]. This structural feature enables the autoencoder to have the advantage of removing irrelevant noise while reconstructing input data. Figure 2 shows the structure of the autoencoder.

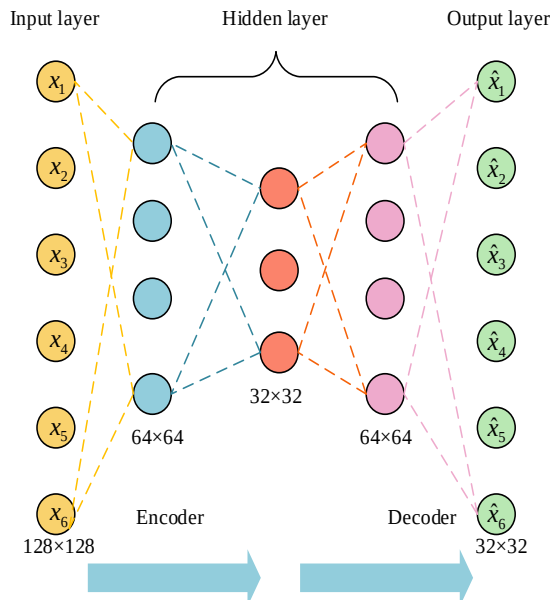


Figure 2: Autoencoder structure

In Figure 2, the autoencoder mainly includes an encoding layer and a decoding layer. When encoding, an autoencoder can convert high-dimensional input data into a low-dimensional latent variable for representation [25]. This latent variable encompasses the key features of the input data and also has a low-level dimension, thereby reducing data complexity and minimizing the need for data storage [26]. During decoding, the autoencoder can remap these low-dimensional latent variable features to

the high-dimensional space of the original input data [27]. The architecture of the autoencoder is as follows. The input layer is $128 \times 128 \times 3$, which represents an image resolution of 128×128 and is an RGB image. The encoder consists of three convolutional layers, with 32, 64, and 128 kernels in the first, second, and third layers, respectively. The kernel size is 3×3 and the stride is 1. The first and second layers of the decoder use 64 and 32 convolution kernels, respectively, with a kernel size of 3×3 and a stride of 1. The output layer consists of one convolutional kernel, with a size of 3×3 and a stride of 1, and uses the Sigmoid activation function. In addition to the Sigmoid activation function, the ReLU activation function is also used in the convolutional layers to handle nonlinear transformations. To prevent over-fitting, Dropout layers are added between the convolutional layers of the encoder, with a dropout probability of 0.2. Equation (5) is the loss function of the autoencoder.

$$L = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2 \quad (5)$$

In equation (5), L means the loss function. N refers to the total datasets. x_i represents the i -th sample in the input dataset. $\hat{x}_i$ is i -th sample in the output dataset. When using traditional self-coding structures for monitoring video abnormal behavior detection, CNN has strong learning and generalization abilities. This can reconstruct input samples in video data that contain abnormal behavior, making it difficult for the model to effectively distinguish between normal and abnormal behavior [28]. Therefore, the study incorporates attention mechanism into the autoencoder for optimization. The attention mechanism can make the model focus more on the parts of the data that contain important information. By introducing variance attention mechanism, autoencoders can adaptively assign higher weights to features with abnormal behavior [29]. In addition, the optimized autoencoder is taken as a generator for GAN to better distinguish between normal and abnormal behavior. The feature block of the attention mechanism is represented by equation (6).

$$\phi(h, w) = \omega * x(h, w) + \text{Attention}(x(h, w)) \quad (6)$$

In equation (6), ϕ represents the feature block sent by the attention mechanism to the convolutional layer for decoding. h and w refer to the rows and columns of the feature map, respectively. $\text{Attention}(x(h, w))$ represents self-attention mechanism. The normalized attention map is represented by equation (7).

$$v(h, w) = \frac{(h, w, d) - \mu}{\sigma} \quad (7)$$

In equation (7), v represents the variance of the normalized attention map. d refers to the depth of the feature map. μ represents the mean of the feature map. σ represents the standard deviation of the feature map. The matching probability of the feature storage module is represented by equation (8).

$$\begin{cases} P = \text{Matching Probability}(F, G) \\ P_t^{k,s} = \frac{\exp((F_s)^T Q_t^k)}{\sum_{s'=1}^S \exp((F_s)^T Q_t^k)} \end{cases} \quad (8)$$

In equation (8), F represents the feature of the feature storage module. G represents the feature output

by the generator. P represents the matching probability. F_s is the item of the feature storage module. Q_t is a feature of the hidden layer. Figure 3 shows the basic framework of GAN.

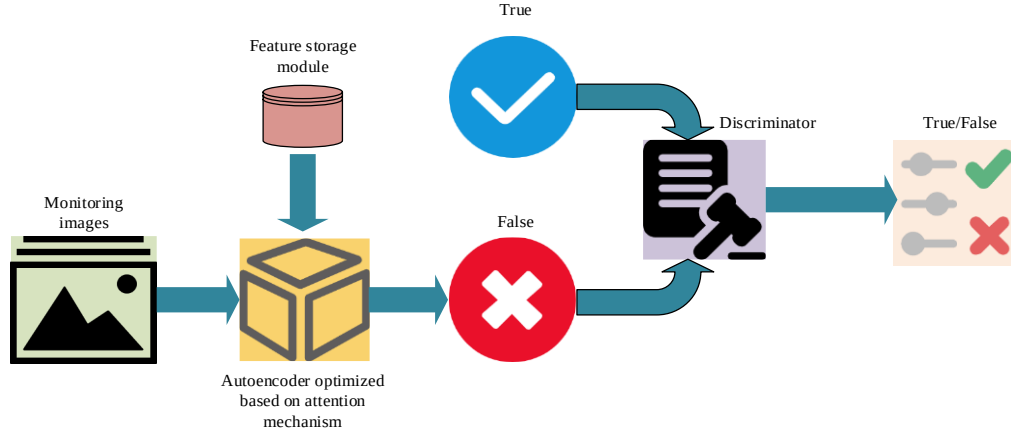


Fig.3 Basic framework of generative adversarial network

In Figure 3, GAN mainly includes two modules: generator and discriminator, which are optimized alternately during the training [30]. The generator is to capture the distribution of real samples and generate outputs similar to real input samples by converting input noise. It can distinguish between real samples and generator generated samples. The discriminator is to distinguish between real samples and fake samples. Using the backpropagation algorithm during training, these two modules alternate for optimization, continuously improving their performance.

In this GAN framework, the generator adopts an optimized deep convolutional autoencoder based on attention mechanism, responsible for feature extraction and reconstruction. Inside the generator, the feature storage module stores key features and participates in generating data during the generation process. The discriminator is responsible for distinguishing between generated data and real data, thereby improving the performance of the generator through adversarial training.

The loss function types of GAN are as follows: The generator loss uses the least squares loss to optimize the generator. The discriminator loss uses adversarial loss to train the discriminator to distinguish between generated images and real images. The training process of this model takes 100 epochs. Within each epoch, the generator and discriminator are trained alternately to ensure training stability. The objective function of GAN is represented by equation (9).

$$\min_G \max_D L(G, D) = E_{a \sim P_a(a)} [\lg D(a)] + E_{z \sim P_z(z)} [1 - \lg D(G(z))] \quad (9)$$

In equation (9), L represents the objective function. G and D refer to generators and

discriminators, respectively. a represents the feature vector of the real monitoring image. z is a noise vector. P_a represents the distribution of real samples. P_z refers to the distribution of noise vectors. E_a represents the expected distribution vector value.

3.2 Clustering algorithm based on mixed multi-input features

The autoencoder network based on deep convolution can handle abnormal behavior detection in ordinary scenes, but there are certain difficulties in accurately detecting abnormal behavior in complex scenes. TSA has limitations such as complex environment, high pedestrian traffic, and multiple foreground targets, all of which can affect detection accuracy. Therefore, a deep convolutional autoencoder network is used to extract feature information from video data based on a Single Shot MultiBox Detector (SSD). Then K-means is introduced to cluster these extracted features. Finally, the distance information of the clustering results and the reconstruction error of the autoencoder are combined to make a comprehensive judgment. This clustering algorithm based on mixed multi-input features helps to enhance algorithm judgment and improve its detection accuracy in complex environments. In the anomaly detection algorithm based on deep convolutional autoencoder network structure, the SSD model is limited to the early stage of object detection as part of video data preprocessing. Its main function is to extract foreground target information for subsequent feature extraction and clustering processing. Figure 4 shows the SSD object detection model.

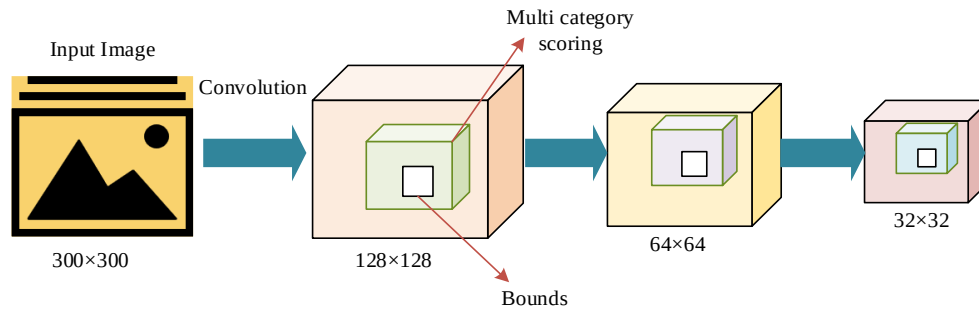


Figure 4: SSD object detection model framework

In Figure 4, this model belongs to a single-stage detection model. It first uses CNN to extract features from sample data, and generates boundary boxes of different sizes on the extracted feature maps. These boundary boxes are used for target classification and prediction. SSD adopts a pyramid structure, allowing for object detection on feature maps with multiple different resolutions. This feature enables the model to have good detection performance for targets of different sizes, thus helping to improve small target detection performance in video surveillance. K-means is an unsupervised learning

algorithm that can measure the similarity between data features by calculating Euclidean distance [31]. Therefore, based on the SSD object detection model and combined with K-means, a clustering algorithm based on mixed multi-input features is designed. This algorithm comprehensively utilizes the feature extraction ability of SSD object detection model and the clustering effect of K-means to achieve more accurate data feature analysis. Figure 5 is a clustering algorithm based on mixed multi-input features.

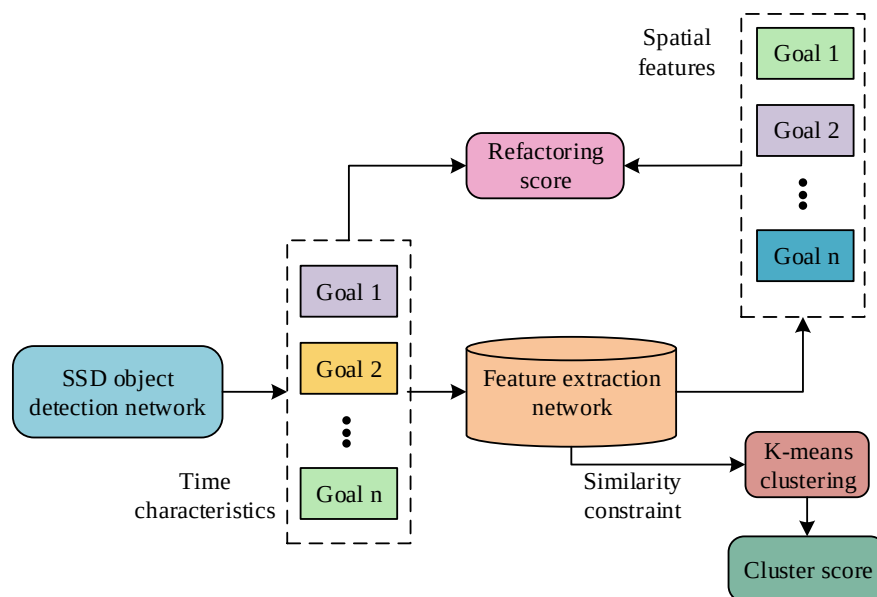


Figure 5: Clustering algorithm structure based on mixed multi-input features

In Figure 5, target n represents the spatiotemporal and temporal features of different foreground targets extracted from video frames. The "reconstruction score" shown in the figure is used to measure the quality of feature reconstruction, which is a key indicator for evaluating abnormal behavior. This score is combined with clustering results to improve the accuracy of detecting abnormal behavior through comprehensive judgment. The clustering algorithm based on mixed multi-input features first utilizes the SSD object detection model to extract foreground target features from multiple input targets. The extractive feature is divided into temporal and spatial features. The time feature information is fed into the feature extraction network as input data for similarity constraints to enhance the

model's recognition ability of time series data. The spatial feature information is constrained by spatial similarity using reconstruction errors to calculate the reconstruction score of information reconstruction quality. Then, the time feature information trained with similarity constraints is input into the clustering module for clustering operations. Based on feature similarity, cluster scores are calculated to evaluate the correlation between targets. In the abnormal behavior detection of scenic area monitoring videos, the motion trajectory of the target object serves as the key basis for determining whether its behavior is abnormal. The SSD configuration is as follows: The resolution of the input image for the SSD model is 300×300 . To handle targets of different sizes, SSD utilizes multiple anchor boxes of different sizes. The

size of the small anchor frame is 32×32 , while the size of the middle anchor frame and the large anchor frame are 64×64 and 128×128 , respectively. The study adopts the k-means++ initialization method to improve the clustering quality and the convergence speed of the algorithm. In the K-means clustering algorithm, based on the characteristics of the dataset, $K=10$ is set to cluster the data into 10 categories. To effectively capture its motion trajectory, the study uses the RGB difference map to analyze the color changes between consecutive frames and map the changes in the motion trajectory. The reconstructed RGB difference map obtained based on behavioral feature transformation is shown in equation (10).

$$\hat{x}_{RGB} = \eta(z_{RGB}; \theta_d^{RGB}) \quad (10)$$

In equation (10), $\hat{x}_{RGB}$ represents the reconstructed RGB difference map obtained based on behavioral feature transformation. η refers to decoding output. z_{RGB} means the behavioral feature generated by the encoder. θ_d^{RGB} is the decoder's parameter set. The loss function during the behavioral feature transformation is represented by equation (11).

$$L_{RGB} = \|x_{RGB} - \hat{x}_{RGB}\|^2 \quad (11)$$

In equation (11), L_{RGB} represents the loss function in converting the RGB difference map into behavioral features. x_{RGB} represents the original RGB difference map of the input. The mathematical expression for clustering score is represented by equation (12).

$$S(r_i) = \sum_{i=1}^N e^{-\alpha \|r_i - c_k\|^2} \quad (12)$$

In equation (12), S represents the clustering score. r_i refers to the i -th feature point extracted from the network. N means the quantity of cluster centers. c_k is the k -th cluster center. α represents the weight of clustering scores. The mathematical expression for the reconstruction score is represented by equation (13).

$$S_m = \|x_m - \hat{x}_m\|_2 \quad (13)$$

In equation (13), S_m represents the reconstruction score. M refers to the quantity of target boxes. The abnormal behavior score is calculated by adding the clustering score and reconstruction score with different weights, represented by equation (14).

$$S(t) = \alpha \sum_{i=1}^{N(t)} S(r_i) + \beta \sum_{m=1}^M S_m \quad (14)$$

In equation (14), S represents the score for abnormal behavior. β refers to the weight of the reconstructed score. $N(t)$ represents the number of features on the t -th frame video image. The score threshold for abnormal behavior scores in this experiment is set to standardized $[0,1]$. The standardized abnormal behavior score is represented by equation (15).

$$S'(t) = \frac{s(t) - s(t)_{\min}}{s(t)_{\max} - s(t)_{\min}} \quad (15)$$

In equation (15), S represents the normalized score of abnormal behaviors. $s(t)_{\max}$ and $s(t)_{\min}$ represent the maximum and minimum scores of abnormal behaviors, respectively. Figure 6 is a clustering algorithm based on mixed multi-input features.

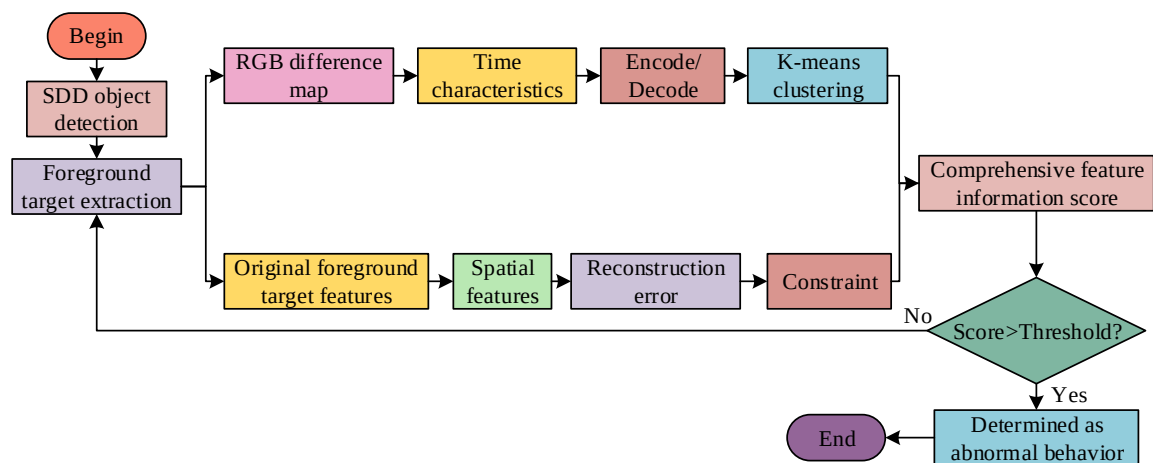


Figure 6: Process of clustering algorithm based on mixed multi-input features

In Figure 6, spatial features are constrained by reconstruction errors. Specifically, the reconstruction error is used to guide the optimization of spatial feature information, thereby improving the sensitivity of the model to spatial features. Unlike directly using reconstruction errors as input features, reconstruction

errors affect spatial features through the output of deep convolutional autoencoders, thereby improving clustering performance and abnormal behavior scoring. In the clustering algorithm based on mixed multi-input features, the SSD object detection model is first used to extract foreground targets from monitoring video images,

thereby reducing the impact of complex environments on extracting effective features. The RGB difference map corresponding to the extracted foreground target is taken as input for CNN, and the temporal feature information of this difference map is extracted. Then, after encoding and decoding by a deep convolutional autoencoder network, these temporal features are constrained to enhance its robustness. The network utilizes reconstruction errors to constrain spatial feature information, ensuring that the output image has effective feature information. K-means is used to classify this constrained feature information and calculate the abnormal behavior score. Finally, the calculated abnormal behavior score is compared with the preset threshold. If the score exceeds the threshold, it is considered abnormal behavior.

4 Verification of abnormal behavior detection in monitoring videos based on deep convolutional autoencoder mixed multi-input feature clustering algorithm

After setting up the experimental environment, the performance of the clustering algorithm was first validated. Then, the effectiveness of the abnormal behavior detection method was verified using methods such as abnormal behavior score detection, ablation experiments, and comparative experiments.

4.1 Experimental environment construction and algorithm performance experiments

To validate the effectiveness of the monitoring video abnormal behavior detection method using multi-input feature clustering, an experimental environment was constructed using the Pytorch framework. A high-performance NVIDIA GeForce RTX 3080 Ti GPU was taken as the cloud host for model training. Meanwhile, an 8-core Intel Xeon CPU was configured for the Windows 10 system to support large-scale data processing. Before conducting the experiment, this input data image was preprocessed and the pixel intensity of monitoring video frames was normalized within $[-1, 1]$. The learning rates of the model generator and discriminator were 0.01 and 0.001, respectively. These datasets used in this experiment are CUHK Avenue and UCSD, which contain monitoring video images collected in natural scenes to distinguish between normal and abnormal behaviors. The CUHK Avenue dataset includes 16 training videos from different scenarios and 21 testing videos, covering various daily activities. These videos include scenes of pedestrians walking normally, while also annotating abnormal behaviors such as running, jumping, and discarding items, providing diverse behaviors in typical urban street environments. The UCSD dataset has two subsets, Ped1 and Ped2, which mainly focus on pedestrian behavior patterns. Ped1 focuses on shooting wider pedestrian areas, while Ped2 focuses more on narrower scenes. These datasets provide video instances of standard walking behavior and various

abnormal behaviors such as cycling and driving. In abnormal behavior detection, key input features include but are not limited to motion trajectories of moving targets, including dynamic parameters such as speed and direction. The appearance features of static parameters such as shape, size, and color extracted using image processing techniques. Table 2 shows the specific experimental environment configuration.

Table 2: Specific experimental environment configurations

Experimental environment	Configuration
Operating system	Windows 10
The Pytorch framework	Pytorch1.8.1
CPU	8 × Intel(R) Xeon(R) CPU E5-2686 v4 @ 2.30GHz
GPU	NVIDIA GeForce RTX 3080 Ti
Memory	64GB
Graphics memory	6G

To verify the performance of DCAMMFCA-based algorithm, a comparison was made between the ordinary clustering algorithm and the anomaly behavior detection algorithm based on the deep convolutional autoencoder. The study sets the training batch to 100 times. Figure 7 presents the accuracy change on the test set. The accuracy based on DCAMMFCA was always higher than that of the other two algorithms. When the training round reached 100, the accuracy of the ordinary clustering algorithm only reached 59.7%. The accuracy of the anomaly detection algorithm based on deep convolutional autoencoder network reached 72.4%. The accuracy of DCAMMFCA reached 89.6%, with an increase of 29.9% and 17.2%, respectively. Therefore, DCAMMFCA, as an independent ensemble algorithm, effectively improves the recognition accuracy of abnormal behavior detection.

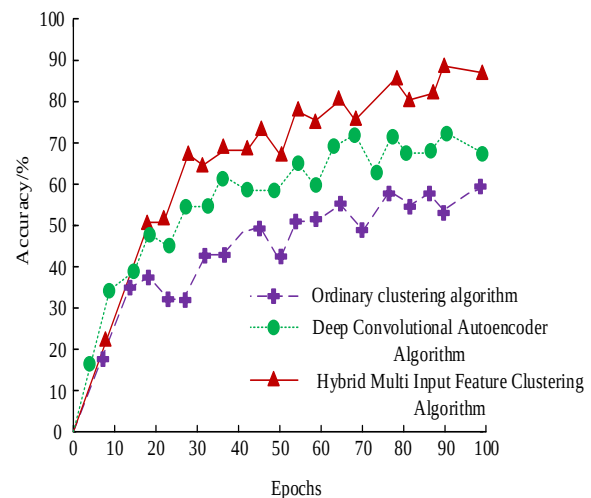


Figure 7: The accuracy variation of different algorithms on the test set

4.2 Performance verification of mixed multi-input feature clustering algorithm

MMFCA, as a fundamental clustering algorithm framework, is used for abnormal behavior detection in video surveillance. The algorithm based on MMFCA has been optimized to DCAMMFCA, which introduces deep convolutional autoencoder and attention mechanism to improve the accuracy of anomaly behavior detection. To verify the performance based on MMFCA, a comparative analysis was conducted between MMFCA and Rough K-Means (RKM), Improved K-Means (IKM), and Fuzzy C-Means (FCM) [32]. The study combined the CUHK Avenue dataset as new data with the UCSD dataset to form an artificial training dataset. Figure 8 shows the distribution of the manually trained dataset, where the data points and clustering centers have undergone preliminary clustering processing. This dataset combines the CUHK Avenue dataset and UCSD dataset for training and testing clustering algorithms. The red data points in the figure represent the clustering of large targets, the blue data points represent the clustering of small targets, the green data points represent the newly added CUHK Avenue dataset data points, and the black data points represent the clustering centers generated by the clustering algorithm. The clustering ratio of large and small targets in artificial datasets is roughly 3:1.

MMFCA was compared with RKM, IKM, and FCM in the artificial training dataset. Figure 9 shows the clustering performance of four algorithms. In Figure 9 (a), RKM failed to identify the newly added data and divided it into small target clusters, resulting in a corresponding shift in the cluster center. In Figure 9 (b), IKM divided the newly added data into large target clusters, causing a shift in the cluster center. In Figure 9 (c), FCM also divided the newly added data into two imbalanced clusters without correctly identifying the new data. In Figure 9 (d), MMFCA effectively identified the newly added data points, and the position of the cluster center was also in the ideal position, showing a high similarity with the distribution of the manually trained dataset. Overall, MMFCA can correctly identify small target data and newly added data, with high recognition accuracy.

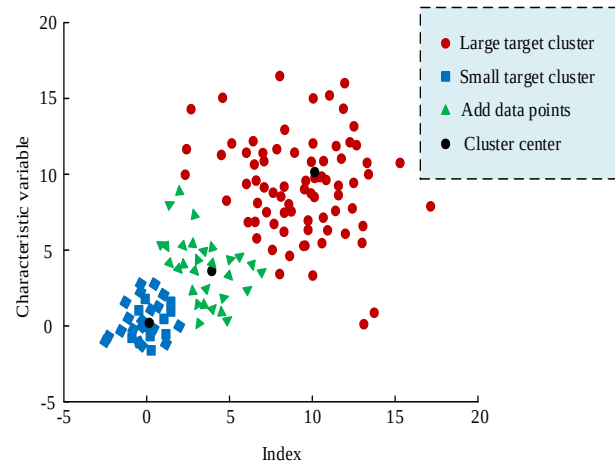


Figure 8: The distribution of artificial training datasets

To further validate the effectiveness of MMFCA, performance indicators such as the Adjusted Rand Index (ARI), Silhouette Coefficient (Sil), and clustering time were compared among these four clustering algorithms. ARI can measure the fitting of clustering algorithms. An ARI close to 1 indicates that its clustering effect is more accurate. Sil can determine the clustering effectiveness. An Sil approaches 1 indicates that the clustering effect is more reasonable. Table 3 presents the performance comparison results of four clustering algorithms. The ARI and Sil of MMFCA were closer to 1, indicating that its clustering effect was closer to the real situation. The ARI was 0.894, which was 142%, 28.8%, and 234% higher than that of RKM, IKM, and FCM, respectively. The Sil of MMFCA was 0.906, which was 116.7%, 50%, and 131.7% higher than that of the other three algorithms, respectively. MMFCA had the shortest clustering time of 0.216s, which was 62.03%, 28.94%, and 27.27% shorter than that of the other three algorithms, respectively. From the F1 score, the research method scored 0.912, while other algorithms all scored over 0.8. In addition, for confusion matrix, the error rate of the research method was the lowest, only at 8%, further proving its superiority. In summary, MMFCA had excellent clustering performance.

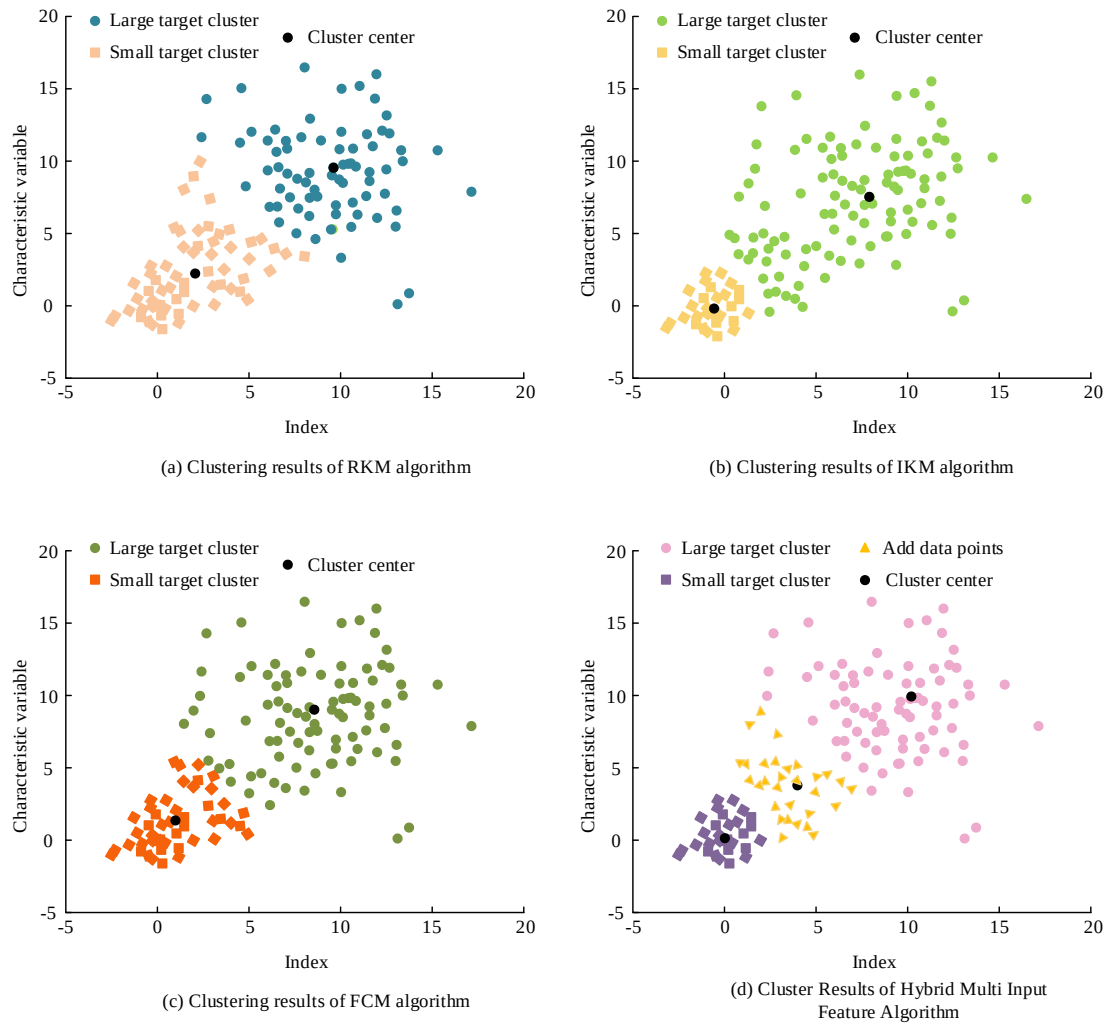


Figure 9: The clustering effect of four algorithms

Table 3: Comparison of performance indicators of four clustering algorithms

Clustering algorithm	ARI	Sil	Cluster time/s	F1 score	Confusion matrix (TP, FN, FP, and TN)
RKM	0.369	0.418	0.569	0.352	(142, 258, 308, and 292)
IKM	0.694	0.604	0.304	0.721	(367, 133, 143, and 357)
FCM	0.267	0.391	0.297	0.288	(121, 279, 302, and 298)
Mixed multi-input feature clustering	0.894	0.906	0.216	0.912	(458, 42, 38, and 462)

4.3 Performance verification of abnormal behavior detection based on deep convolutional autoencoder mixed multi-input feature clustering algorithm

To validate the abnormal behavior detection performance, this study compared this detection method with abnormal behavior detection methods such as Adam, MDT, SF, and SRC [33]. To further validate the stability of the model performance, a 95% confidence interval was added when calculating the ROC curve. All AUC values were the average based on five-fold cross-validation. Figure 10 presents the Receiver Operating Characteristic (ROC) curves using different abnormal behavior detection

methods. In Figure 10 (a), on the CUHK Avenue data, the AUC value of the abnormal behavior detection method based on DCAMMFCA was 91.9%, which was 41.8%, 13.0%, 19.5%, and 3.1% higher than the AUC values of the Adam, MDT, SF, and SRC anomaly detection methods, respectively. In Figure 10 (b), on the UCSD dataset, the AUC value of the abnormal behavior detection method based on DCAMMFCA was 94.7%, which was 48.6%, 10.6%, 38.2%, and 4.7% higher than that of the other four abnormal behavior detection methods, respectively. Overall, the abnormal behavior detection method based on DCAMMFCA had high detection accuracy.

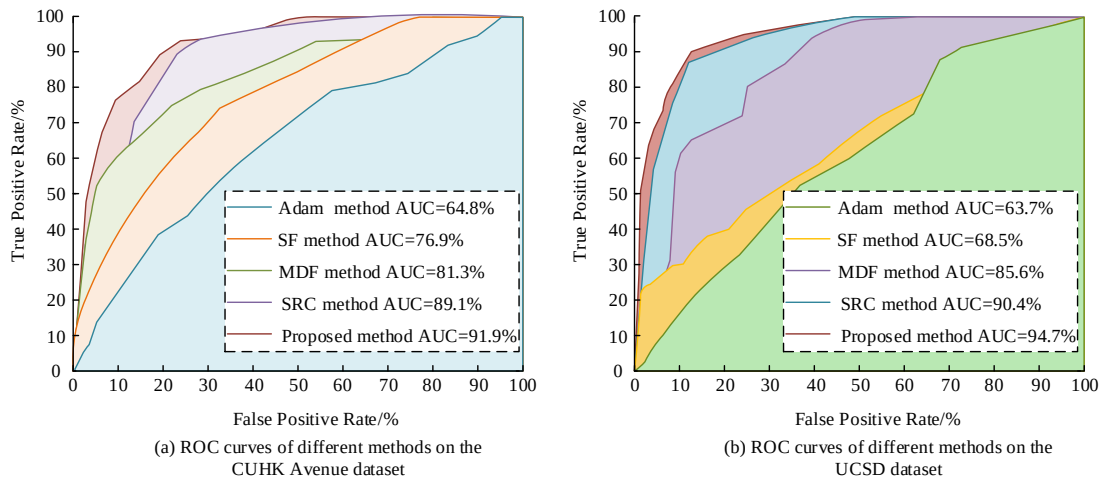


Figure 10: The ROC curve of different abnormal behavior detection methods

To observe the performance of different detection methods more intuitively, the performance of different abnormal behavior detection methods on different datasets was compared. Table 4 shows the comparison results of performance indicators for different abnormal behavior detection methods. The abnormal behavior detection method based on DCAMMFCA achieved better performance on different datasets. On the CUHK Avenue

dataset, the precision, recall, and error rate of this method were 94.5%, 95.6%, and 12.6%, respectively. On the UCSD dataset, the precision, recall, and error rate of this method were 95.2%, 94.8%, and 10.9%, respectively. Overall, the detection precision of the abnormal behavior detection method based on DCAMMFCA was superior to that of the other four abnormal behavior detection methods.

Table 4: Comparison of different abnormal behavior detection methods' performance indicators

Test method	CUHK Avenue dataset			UCSD dataset		
	Precision/%	Recall/%	Error rate/%	Precision/%	Recall/%	Error rate/%
Adam	53.4	64.2	39.1	55.3	66.3	41.9
MDT	73.1	72.6	25.6	70.6	74.0	25.2
SF	60.4	68.6	30.5	61.6	69.1	41.7
SRC	88.4	82.1	19.6	87.5	83.1	15.6
Proposed method	94.5	95.6	12.6	95.2	94.8	10.9

4.4 Abnormal behavior score detection and ablation experiment

To validate the abnormal behavior detection effectiveness in practical applications, this study compared it with different abnormal behavior detection methods for abnormal behavior score detection, as shown in Figure 11. The abnormal behavior score was calculated through K-means clustering and reconstruction score. The abnormal behavior score of each video frame was compared with the normal behavior score, from which the difference value between the frame and the normal behavior score was calculated. In actual monitoring videos, the difference in abnormal scores of the abnormal behavior detection method based on DCAMMFCA was 0.297, which was 34.38%, 16.93%, 22.22%, and 16.01% higher than the difference in abnormal scores of abnormal behavior detection methods such as Adam, MDT, SF, and SRC, respectively. The abnormal behavior detection method based on DCAMMFCA had high accuracy in identifying abnormal behaviors.

The training set in the experiment contains 5,000 samples, and the testing set contains 1,000 samples. The training and testing sets were randomly selected from the UCSD

and Avenue datasets, ensuring the diversity and representativeness of the experimental data. To further verify the role of different modules in the abnormal behavior detection method, the study added each module to the network for ablation experiments. "√" indicates the presence of the module, and "/" indicates the absence. Table 5 shows the ablation experiment. After optimizing the autoencoder using attention mechanism, the accuracy improved by 3.2%. After introducing GAN, the accuracy improved by 13.5%. When the algorithm was added to the SSD object detection model, the accuracy improved by 8.6%. After adding K-means, the accuracy improved by 5.8%. In addition, to evaluate the stability of the model, a five-fold cross-validation was conducted to calculate the standard deviation of the results. After adding various modules, the standard deviation gradually decreased from ± 1.5 to ± 0.8 , demonstrating the stability of the model under different experimental configurations. In summary, the added modules brought benefits to the abnormal behavior detection, indicating that the proposed method effectively improved the abnormal behavior detection performance.

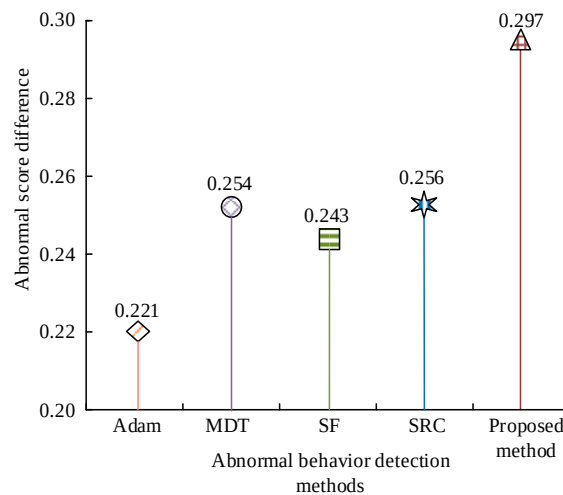


Figure 11: Comparison of abnormal behavior score detection

Table 5: Ablation experiment

Autoencoder	Attention mechanism	GAN	SSD object detection	K-means	Accuracy/%	Standard deviation
√	/	/	/	/	59.4	±1.5
√	√	/	/	/	62.6	±1.2
√	√	√	/	/	76.1	±1.0
√	√	√	√	/	84.7	±0.9
√	√	√	√	√	90.5	±0.8

5 Discussion

To improve the detection accuracy of tourist attraction monitoring, the DCAMMFCA method for detecting abnormal behavior in monitoring videos of tourist attractions is proposed. Compared with the existing state-of-the-art methods for detecting abnormal behaviors [32–33], the research method has shown significant advantages in multiple indicators. The method based on DCAMMFCA showed significantly higher detection accuracy on both the CUHK Avenue and UCSD datasets. For example, on the CUHK Avenue dataset, the accuracy of DCAMMFCA was 94.5%, which was approximately 41.8%, 13.0%, 19.5%, and 3.1% higher than other methods. This is because DCAMMFCA combined with GAN and attention mechanism can significantly improve the performance of abnormal behavior detection, especially in complex environments and small object detection. In terms of computational efficiency, the DCAMMFCA method had an average computation time of 0.216 seconds on the CUHK Avenue dataset, which was significantly lower than other algorithms. Especially compared with clustering algorithms such as RKM, IKM, and FCM, the efficiency was improved by 62.03%. This method introduces GAN and attention mechanism, which can maintain high robustness in constantly changing environments and have high computational efficiency, making it suitable for real-time monitoring applications. In terms of the reliability of abnormal behavior classification, this method achieved a recognition accuracy of 89.6% for small targets, and the clustering

effect was highly consistent with the real situation. The ARI and Sil were 0.894 and 0.906, respectively, close to 1, indicating the superiority of clustering effect.

6 Conclusion

To improve the public safety of TSA, a self-encoder structure GAN optimized by attention mechanism was built, and the SSD object detection model combined with multi-input feature clustering algorithm was used to improve the accuracy of small object detection. These results confirmed that the accuracy of DCAMMFCA reached 89.6%. The ARI and Sil reached 0.894 and 0.906, respectively, which were close to 1, indicating that its clustering effect was close to the real situation. In terms of computation time, MMFCA took 0.216 seconds, which was 62.03%, 28.94%, and 27.27% shorter than that of RKM, IKM, and FCM, respectively. On the datasets CUHK Avenue and UCSD, the AUC values of the abnormal behavior detection method based on DCAMMFCA reached 91.9% and 94.7%, respectively, far higher than that of the other four behavioral anomaly detection methods. On the CUHK Avenue dataset, the precision, recall, and error rate of this method were 94.5%, 95.6%, and 12.6%, respectively. On the UCSD dataset, the precision, recall, and error rate of this method were 95.2%, 94.8%, and 10.9%, respectively. The abnormal score difference was 0.297, which was 25.58%, 14.47%, 18.18%, and 13.80% higher than that of Adam, MDT, SF, and SRC, respectively. In summary, the research on TSA monitoring video abnormal behavior detection based on MMFCA had effectively improved the

accuracy of abnormal behavior detection.

However, there are still some limitations in the research, such as the lack of interpretability layers, failure to test on real TSA datasets, domain adaptation/generalization issues, and insufficient evaluation under adversarial and occlusion conditions. In response to these limitations, future work can further classify the types of abnormal behaviors detected to take different measures to deal with different types of abnormal behaviors. In addition, to adapt to different scenarios and data distributions, transfer learning methods can be explored in the future to quickly adapt to new monitoring environments with a small amount of annotated data. In addition, methods for multi-modal data fusion can be explored, such as combining the thermal imaging and RGB images to improve the accuracy and robustness of abnormal behavior detection.

References

- [1] Lentzas A and Vrakas D. Non-intrusive human activity recognition and abnormal behavior detection on elderly people: A review. *Artif. Intell. Rev.*, vol. 53, no. 3, pp. 1975–2021, Mar. 2020. DOI: 10.1007/s10462-019-09724-5
- [2] Alafif T, Hadi A, Allahyani M, Alzahrani B, Alhothali A, Alotaibi R, and Barnawi A. Hybrid classifiers for spatio-temporal abnormal behavior detection, tracking, and recognition in massive Hajj crowds. *Electron.*, vol. 12, no. 5, pp. 1165–1173, February. 2023. DOI: 10.3390/electronics12051165
- [3] Roka S and Diwakar M. CViT: A convolution vision transformer for video abnormal behavior detection and localization. *SN Comput. Sci.*, vol. 4, no. 6, pp. 829–834, October. 2023. DOI: 10.1007/s42979-023-02294-y
- [4] Chen N, Man Y, and Sun Y. Abnormal cockpit pilot driving behavior detection using YOLOv4 fused attention mechanism. *Electron.*, vol. 11, no. 16, pp. 2538–2541, August. 2022. DOI: 10.3390/electronics11162538
- [5] Wang B, Jiang X, Dong Z, and Li J. Behavioral parameter field for human abnormal behavior recognition in low-resolution thermal imaging video. *Appl. Sci.*, vol. 12, no. 1, pp. 402–415, December. 2021. DOI: 10.3390/app12010402
- [6] Jenssen R and Roverso D. Intelligent monitoring and inspection of power line components powered by UAVs and deep learning. *IEEE Power Energy Technol. Syst. J.*, vol. 6, no. 1, pp. 11–21, January. 2019. DOI: 10.1109/JPETS.2018.2881429
- [7] Yousefi H, Su H M, Imani S M, Alkhaldi K, Filipe C D M, and Didar T F. Intelligent food packaging: A review of smart sensing technologies for monitoring food quality. *ACS Sens.*, vol. 4, no. 4, pp. 808–821, March. 2019. DOI: 10.1021/acssensors.9b00440
- [8] Pimenov D Y, Bustillo A, Wojciechowski S, Sharma V S, Gupta M K, and Kuntoğlu M. Artificial intelligence systems for tool condition monitoring in machining: Analysis and critical review. *J. Intell. Manuf.*, vol. 34, no. 5, pp. 2079–2121, March. 2023. DOI: 10.1007/s10845-022-01923-2
- [9] Liu B. Based on intelligent advertising recommendation and abnormal advertising monitoring system in the field of machine learning. *International Journal of Computer Science and Information Technology*. 2023 Dec, vol. 1, no. 1, pp. 17–23. DOI:10.62051/ijcsit.v1n1.03
- [10] Mattera G, Nele L, Paoletta D. Monitoring and control the wire arc additive manufacturing process using artificial intelligence techniques: a review. *Journal of Intelligent Manufacturing*. 2024 Feb, vol. 35, no. 2, pp. 467–97. DOI:10.1007/s10845-023-02085-5
- [11] Aldhamari A, Sudirman R, and Mahmood N H. Abnormal behavior detection using sparse representations through sequential generalization of k-means. *Turk. J. Electr. Eng. Comput. Sci.*, vol. 29, no. 1, pp. 152–168, June. 2021. DOI: 10.3906/elk-1904-187
- [12] Hu J, Zhang X, and Maybank S. Abnormal driving detection with normalized driving behavior data: A deep learning approach. *IEEE Trans. Veh. Technol.*, vol. 69, no. 7, pp. 6943–6951, July. 2020. DOI: 10.1109/TVT.2020.2993247
- [13] Feizi A. Hierarchical detection of abnormal behaviors in video surveillance through modeling normal behaviors based on AUC maximization. *Soft Comput.*, vol. 24, no. 14, pp. 10401–10413, July. 2020. DOI: 10.1007/s00500-019-04544-9
- [14] Zhang H and Zhou W. A two-stage virtual machine abnormal behavior-based anomaly detection mechanism. *Cluster Comput.*, vol. 25, no. 1, pp. 203–214, February. 2022. DOI: 10.1007/s10586-021-03385-2
- [15] Gao H, Zhou L, Kim JY, Li Y, Huang W. Applying probabilistic model checking to the behavior guidance and abnormality detection for A-MCI patients under wireless sensor network. *ACM Transactions on Sensor Networks*. 2023 Mar, vol. 19, no. 3, pp. 1–24. DOI:10.1145/3499426
- [16] Tan R, Tan Q, Zhang Q, Zhang P, Li Z. Ethereum fraud behavior detection based on graph neural networks. *Computing*. 2023 Oct, vol. 105, no.10, pp. 2143–70. DOI:10.1007/s00607-023-01177-7
- [17] Popescu D, Stoican F, Stamatescu G, Ichim L, and Dragana C. Advanced UAV–WSN system for intelligent monitoring in precision agriculture. *Sensors*, vol. 20, no. 3, pp. 817–823, February. 2020. DOI: 10.3390/s20030817
- [18] Collins G S, Moons K G M. Reporting of artificial intelligence prediction models. *The Lancet*, vol. 393, no. 10181, pp. 1577–1579, October. 2019. DOI: 10.1016/S0140-6736(19)30037-6.
- [19] Huang Z, Xu Y, Cheng Y, Xue M, Deng M, Jaffrezic-Renault N, and Guo Z. Recent advances in skin-like wearable sensors: Sensor design, health monitoring, and intelligent auxiliary. *Sensors & Diagnostics*, vol. 1, no. 4, pp. 686–708, July. 2022. DOI: 10.1039/D2SD00009H.

- [20] Hashimoto D A, Witkowski E, Gao L, Meireles O, and Rosman G. Artificial intelligence in anesthesiology: current techniques, clinical applications, and limitations. *Anesthesiology*, vol. 132, no. 2, pp. 379-394, February. 2020. DOI: 10.1097/ALN.0000000000002960
- [21] Shi Q, Zhang Z, He T, Sun Z, Wang B, Feng Y, ... and Lee C. Deep learning enabled smart mats as a scalable floor monitoring system. *Nat. Commun.*, vol. 11, no. 1, pp. 4609-4624, February. 2020. DOI: 10.1038/s41467-020-18471-1.
- [22] Vaishya R, Javaid M, Khan I H, and Haleem A. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 14, no. 4, pp. 337-339, July. 2020. DOI: 10.1016/j.dsx.2020.04.012.
- [23] Jiao J, Zhao M, Lin J, and Ding C. Deep coupled dense convolutional network with complementary data for intelligent fault diagnosis. *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9858-9867, February. 2019. DOI: 10.1109/TIE.2019.2894765.
- [24] Rahaman A, Islam M M, Islam M R, Sadi M S, and Nooruddin S. Development IoT Based Smart Health Monitoring Systems: A Review. *Rev. d'Intelligence Artif.*, vol. 33, no. 6, pp. 435-440, March. 2019. DOI: 10.18280/ria.330601.
- [25] Motwani A, Shukla P K, Pawar M. Novel framework based on deep learning and cloud analytics for smart patient monitoring and recommendation (SPMR). *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 5, pp. 5565-5580, February. 2023. DOI: 10.1007/s12652-022-03960-2.
- [26] Fatema A, Poondla S, Mishra R B, and Hussain A M. A low-cost pressure sensor matrix for activity monitoring in stroke patients using artificial intelligence. *IEEE Sens. J.*, vol. 21, no. 7, pp. 9546-9552, March. 2021. DOI: 10.1109/JSEN.2021.3054637.
- [27] Verma K K, Singh B M, Dixit A. A review of supervised and unsupervised machine learning techniques for suspicious behavior recognition in intelligent surveillance system. *Int. J. Inf. Technol.*, vol. 14, no. 1, pp. 397-410, October. 2022. DOI: 10.1007/s41870-020-00519-8.
- [28] Yao D, Wen M, Liang X, Fu Z, Zhang K, and Yang B. Energy theft detection with energy privacy preservation in the smart grid. *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7659-7669, March. 2019. DOI: 10.1109/JIOT.2019.2915041.
- [29] Mills M C, Rahal C. The GWAS Diversity Monitor tracks diversity by disease in real time. *Nat. Genet.*, vol. 52, no. 3, pp. 242-243, March. 2020. DOI: 10.1038/s41588-020-0580-y
- [30] Santhosh K K, Dogra D P, Roy P P. Anomaly detection in road traffic using visual surveillance: A survey. *ACM Comput. Surv. (CSUR)*, vol. 53, no. 6, Article 1, pp. 1-26, February. 2020. DOI: 10.1145/3397271.
- [31] Yuan P, Fan C, Zhang C. YOLOv5s-MEE: A YOLOv5-based Algorithm for Abnormal Behavior Detection in Central Control Room. *Information Technology and Control*. 2024 Mar, vol. 53, no. 1, pp. 220-36.
- [32] Wang Y, Liu Y, Feng W, and Zeng S. Waste Haven Transfer and Poverty-Environment Trap: Evidence from EU. *Green Low-Carbon Econ.*, vol. 1, no. 1, pp. 41-49, February. 2023. DOI: 10.47852/bonviewGLCE3202668.
- [33] Wellendorf A, Tichelmann P, Uhl J. Performance Analysis of a Dynamic Test Bench Based on a Linear Direct Drive. *Arch. Adv. Eng. Sci.*, vol. 1, no. 1, pp. 55-62, June. 2023. DOI: 10.47852/bonviewAAES3202902

