Detection of Abnormal Human Behavior in Video Images based on a Hybrid Approach

BAI Ya-meng*, WANG Yang, WU Shen-shen
School of Information Engineering, Jiaozuo University, Jiaozuo 412000, China

***Abstract*—The analysis of human movement has attracted the attention of many scholars of various disciplines today. The purpose of such systems is to perceive human behavior from a sequence of video images. They monitor the population to find common properties among pedestrians on the scene. In video surveillance, the main purpose of detecting specific or malicious events is to help security personnel. Different methods have been used to detect human behavior from images. This paper has used an efficient computational algorithm for detecting anomalies in video images based on the combined approach of the differential evolution algorithm and cellular neural network. In this method, the input image's gray-level image is first generated. Because it may be possible to identify several large areas in the image after the threshold, the largest white area is selected as the target area. The images are then used to remove noise, smooth the image, and fade the morphology. The results showed that the proposed method has higher speed and accuracy than other methods. The advantage of the algorithm is that it has a runtime of three seconds on a home computer, and the average sensitivity criterion is 98.6% (97.2%).**

Keywords—Cellular neural network; detection of abnormalities; differential evolution algorithm; video images

I. Introduction

The analysis and investigation of human movements have attracted the attention of many researchers in different fields. Such systems aim to understand human behavior from a sequence of video images [1, 2]. Security cameras are installed in many organizations that are important for security and can monitor and record the situation of a place at all times. Human behavior detection systems are also one of the most important needs of surveillance systems in the sequence of video images captured by surveillance cameras [3, 4, 5]. Usually, all automatic detection or monitoring systems limit the detection mechanisms of the desired objects in the field of view to the tracking and activities of objects in the same image areas. Crowd monitoring is used to identify general characteristics among passers-by in a scene [6, 7, 8, 9]. These properties include population size, density, growth rate, traffic patterns, and the detection of abnormal events. In video surveillance, the main goal is to detect specific or abnormal events to help security personnel. Surveillance systems in use today rely on the performance of a human operator. They are expected to view many screens recorded by different cameras, often simultaneously. Generally, classification is one of the most important and difficult steps in image processing [10, 11]. Various methods have been used to recognize human behavior from images, but some of these methods do not show optimal performance in some situations, and in some cases, the methods used are either not able to detect people under certain conditions or do not have high accuracy [12, 13, 14, 15, 16]. Another problem with several methods is that they require a powerful processing unit to achieve the desired detection speed. Among the significant problems, when there is a need to model busy and complex scenes due to occlusion, the use of background reduction to classify motion areas and the trajectory method for feature extraction fail [17, 18, 19]. Some of these problems can be seen in Fig. 1. As seen in Fig. 1, people have much variety in the shape and color of their clothes. Also, different lighting conditions can be seen in the images. There is also a partial or total overlap of the person's image with other objects and people [20, 21, 22, 23]. The superiority of cellular neural networks in low-level image processing over common digital image processing systems is due to two features of this network. These features include parallel processing of the analog image signal and cross-sectional or local connections among network cells, making it easy to build cellular neural networks. In this research, abnormalities will be detected from video images. For this, the cellular neural network method will be used. Since this network needs optimized parameters, the differential evolution algorithm will be used to optimize the features for training. This algorithm is similar to the genetic algorithm, except that changes have been made in the operations of mutation and combination. The differential evolution algorithm will use the training data set to calculate the fitness function for the produced chromosome. Based on this, the template generated by the differential evolution algorithm is coded. Then, one of the two main images from the database is randomly selected. After producing a suitable template for the cellular neural network, all the images in the database will be given to the network for segmentation. Recognizing moving objects and people in the dark or in environments with insufficient light is one of the limitations of this research.

A. Motivation and Contributions

Detecting and tracking moving people (such as the abnormality of a person in several consecutive images of a scene) has many challenges, including the great variety of people in terms of shape, color, and clothing. It is also possible that, according to the angle of placement of people in front of the camera, the person's appearance may change, or a person may overlap with another person or other objects in the scene. For this reason, it isn't easy to detect and recognize people in crowded environments. Therefore, local motion information should be extracted at the pixel level using spatial and temporal detectors and the optical flow characteristics of shape, color, and clothing to extract features in such scenes. It

*Corresponding Author.

is also possible that, according to the angle of placement of people in front of the camera, the person's appearance may change, or a person may overlap with another person or other objects in the scene.

For this reason, it isn't easy to detect and recognize people in crowded environments. Therefore, local motion information should be extracted at the pixel level using spatial and temporal detectors and optical flow characteristics to extract features in such scenes. It will be possible to view the video image. The contributions of the authors in the article are as follows:

- Using a cellular neural network and a differential evolution algorithm to improve the accuracy of finding things that aren't normal.
- Increasing classification sensitivity in abnormality detection based on video images.

The continuation of the paper is configured as follows: The second section discusses previous work in detecting abnormal human behavior in video images. In the third part, the proposed method is given in detail. In the fourth part, the results of the proposed method are compared with similar methods. Finally, general conclusions and suggestions for future work are stated in the fifth part.

Fig. 1. Complications in Identifying People [1].

II. RELATED WORKS

This section introduces some methods to detect human behavior abnormalities in video images. It is presented a method for detecting local and series anomalies in crowded scenes using k=24 bucket representation with sparse (MHOF). In this method, the multi-comparison histogram of the optical flow on three scales is used to accurately store the movement direction information and the movement energy information to express the characteristics [24]. Another support vector regression machine (SVR) method has been used for pedestrian detection. In this paper, to extract the feature, the features of the Harr wavelet and directional histogram have been used, and in the classification section, in addition to using SVR, the classifier K-nearest neighbor (KNN) has also been used. The results of the experiments have shown the high accuracy of the algorithm presented in this paper [25]. Researchers used fuzzy logic to detect anomalies in human behavior. In this paper, to determine which rules are suitable, the effectiveness of each rule was evaluated based on the speed and environmental conditions of the pedestrian. This method is very time-consuming due to the type of search that requires the search window to be moved pixel-by-pixel in both horizontal and vertical directions and is not suitable for real-time or live applications [26]. Also, researchers have used graph formulation to diagnose human behavioral abnormalities. In this method, a model of the appearance and dynamics of the scene is created, and the features of the pixel level are used. Each designated location defines the center of the video cell, from which the edges and vertices of the graph are extracted. A tree is trained during the training phase. The probability of each of the obtained spatio-temporal volumes and the representation of its salient points are generated. An anomaly is where the probability is less than a threshold. The results showed that using graphs provided complete coverage of the problem space. However, it takes a long time to process and calculate [27]. In this research, optical flow and gradient have been used to detect anomalies. As a result, among these methods, a bag of words is a suitable method for distinguishing abnormal combinations from normal events, especially in complex scenes that include different classes of objects and events. Like anomalies in traffic scenes, a method for rapid anomaly detection for surveillance systems, with real-time capability and reliable detection and localization, is proposed. Each event includes a set of spatio-temporal volumes. This method creates a model of normative behavior by initializing the algorithm with a small number of video sequences, then seeks to detect anomalies and gradually updates the model [28]. Researchers tried to speed up pedestrian detection by using a neural network. In this method, they were recognized based on the rescaling method in deep learning and by sharing features across the model. The results of this method have been better than the normal neural network in terms of processing speed [29]. In order to build an advanced driver assistance system (ADAS), researchers introduced a method based on fuzzy clustering and convolutional neural networks. In this paper, the candidate areas are revealed first. Then a quick method is used to extract the features. In order to extract the candidate areas, the fuzzy clustering method is used. After the features are extracted, to check whether there are pedestrians in those candidate areas or not, neural networks have been used. This type of neural network is made by imposing restrictions on weights and how to connect the neurons of a standard neural network and creates a network suitable for processing data with temporal or spatial distribution [30]. In another study, pedestrians were recognized based on gradient and texture features. This method used principal component analysis (PCA) to reduce the dimensions of textural and rotational features. Experiments on two different data sets showed the better efficiency of this approach compared to similar methods [31]. A robust method for human detection in underground image sequences that works based on learning systems was presented. An SVM support vector machine classifier is used for detection, so that after the preprocessing process, the patterns of the image in which human presence is likely are extracted and fed into the SVM. The OSU pedestrian database has been used to learn and test the proposed algorithm. The results of the implementation of the presented algorithm on this database show the efficiency and appropriateness of the algorithm [32]. They presented another method to remove the background in the images of the outside environment that does not have a fixed background. There are things like the movement of leaves and moving flags in them. Their main idea is that the pixels of the neighboring blocks, which are assumed to be the background, have similar changes over time and that the same patterns can be extracted from them. Although this argument is correct for the blocks extracted from an object in the background, it does not seem correct for the blocks that fall on the boundary of the objects separately due to different changes on both sides of the boundary. They are used to support pedestrian detection based on underground images. This method was performed on a dataset with 18 images, and the detection accuracy of this method was 98.11% [33]. The study's [34] objective is to train the detector to identify a human action even when only a partial action example is present. To achieve this, a hybrid approach based on the fuzzy Bandler subproduct and the Kohut triangle is proposed in this work that combines the benefits of computer vision and fuzzy set theory. Building a frame-by-frame membership function for each potential motion type is novel. When a predetermined threshold is appropriately reached, detection is started. The benefits and efficacy of the suggested method are demonstrated by experimental results on a dataset that is publicly accessible. Source [35] offers a novel method for identifying and detecting human behavior. This work's major goal is to show how to identify abnormal behavior in a constrained domain. The suggested method automatically detects people in regular films, or video surveillance feeds. When a person is considered in a frame, the action is recognized, and the human position is approximated. After that, the behavior is categorized as normal or abnormal. The models included in the proposed work include SSD MOBILE NET for human action identification and the FAST R-CNN model for human detection in a video frame or image. Everyone needs to feel safe and secure at all times in our century. Many nations have purchased and implemented surveillance systems to secure their environment using high-definition CCTV cameras. As a result, automated CCTV surveillance systems can operate as security providers by spotting suspicious human behavior or intruder activity using CCTV footage. This is a difficult task. Models of suspicious behavior that only partially employ machine learning techniques can be used to do this. It has been discovered that several of the prior methods used techniques from deep learning, machine learning, IoT, and fuzzy logic. The proposed suspicious behavior detection model (SHBDM) efficiently extracts picture features using a CNN model pre-trained on the ImageNet dataset known as Inception V3 (VGG-16). Python was used to create this system using an open-source framework. Compared to VGG-16 + LSTM and a straightforward CNN model, the precision accuracy of the system using Inception V3 + LSTM was enhanced by 88.8% [36]. Table I is an overview of the detection of abnormal human behavior in video images, whose problems and solutions have been studied in previous works.

TABLE I. OVERVIEW OF DETECTION OF ABNORMAL HUMAN BEHAVIOR IN VIDEO IMAGES

Work	Method	Solution
[37]	large numbers of people who never stop moving. The vertical arms and legs reaching out from the person distort the blobs' alignment.	Extrinsic appearance (geometry) and intrinsic appearance are used to parameterize the current mark
[38]	High volume video data, for example, has intrinsic non-linearity, spatial localization of patterns, and noise, all of which must be dealt with by the complex interplay between Multiobject and the uncontrolled scene.	Demonstrate a subspace detector that learns non-linearly
[39]	Use of Deep Metric Learning to Take Advantage of Label Correlation in a Transferred-Mediface Sample	The current bottom-up method focuses on a person's short-term behavior and trajectory.
[40]	It's challenging to make direct comparisons between motion and appearance representations, which are usually adapted to a certain scene domain.	Show a unified (temporal and spatial) model of form and movement (Mixture of dynamic texture)
[41]	Detection in a crowd scene may be impacted by factors like occlusion and fluctuating lighting.	The foreground was extracted using a double filtering technique, and the noise was diminished using a media filter.
[42]	Crowds make it hard to identify and follow individual people. When faced with unexpected situations, learning, adaptive, and incremental approaches are not effective.	Exhibit a tiered strategy encompassing low, medium, and high
[43]	Due to the intricate temporal dynamics and interactions among the behavior of numerous objects, anomalies tend to be subtle.	Implementing a Staggering Arrangement of Evolving Bayesian Networks
[44]	Analysis of crowd behavior is challenging. Complex behaviors such as line formation, laminar and turbulent flow, arching and clogging at exists, jamming around barriers, and panic have been seen in the movements of crowds.	A grid of particles was placed on top of the image, and the particles were advected using the average space-time optical flow.
[45]	Normal behavior in a video scenario is difficult to predict, necessitating intensive labeling and training. Due to the high number of objects in a crowd scene, it is challenging to precisely monitor individual objects.	Using a patch-based local motion representation, introduce the novel idea of contextual anomaly into crowd analysis.
[46]	In occlusions, the mobility of the people and things in the scene is exceedingly erratic. In a film with that many people, dissecting the behavior of each individual is an arduous process.	Exhibit a 3D Gaussian gradient in space and time that is rich, non-uniform, and localized.

III. Suggested Approach

The aim of this research is to increase the accuracy of abnormality detection using a cellular neural network based on video images. Observations show that, when there is a need to model busy and complex scenes, due to the presence of occlusion, the use of background subtraction for zoning motion regions and the trajectory method for feature extraction fail. So, to get features out of these kinds of scenes, local motion information should be extracted at the pixel level with the help of spatial and temporal detectors and optical flow features. The proposed detection system provides an improved method for feature selection using the differential evolution algorithm in the image using a cellular neural network. Fig. 2 shows the block diagram of the proposed approach.

B. Reading the Image

The first step in the proposed approach is to pre-process the video image to obtain the anomaly region. Since the changes in the intensity of the gray levels in the images prevent the accurate detection of the area related to the abnormality, therefore, in the proposed approach, the lighting improvement technique is used to linearly map the intensity values in the image to new values that are normally distributed in the range of 0 to 255. are used. Also, to reduce existing noise and improve the image, a common median filter has been used.

C. Noise Reduction

The histogram is balanced by increasing the intensity values with the highest frequency. This method is useful in images where both background and foreground are dark, or both are light. In particular, this method can lead to better details in images produced under or overexposed to X-rays. In this paper, a median filter is used to reduce its noise. This filter is a non-linear digital image filtering method that considers each pixel and examines its nearest neighbors to determine a median value. Then the value of the pixels is replaced with the determined median value.

D. Increase Contrast

The first step in pre-processing is to use Gaussian function to increase the contrast in the original image. It is used before identifying the edges. Gaussian function in two-dimensional space is shown as Eq. (1) [4].

$$G_\sigma = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{i^2+j^2}{2\pi\sigma^2}\right)} \text{ ، } i = 1،2،..،n، \ j = 1،2،..،m \qquad (1)$$

Gaussian filter plays an important role in many feature extraction methods in such a way that the implementation of its function is equivalent to filtering the image using the Gaussian kernel and non-uniform sampling of the image.

E. Feature Extraction

The shape of objects is one of the important features in detecting anomalies in video images. Usually, anomalies with similar shapes belong to several classes. Therefore, in this research, most of the extracted features are related to the shape of the abnormality.

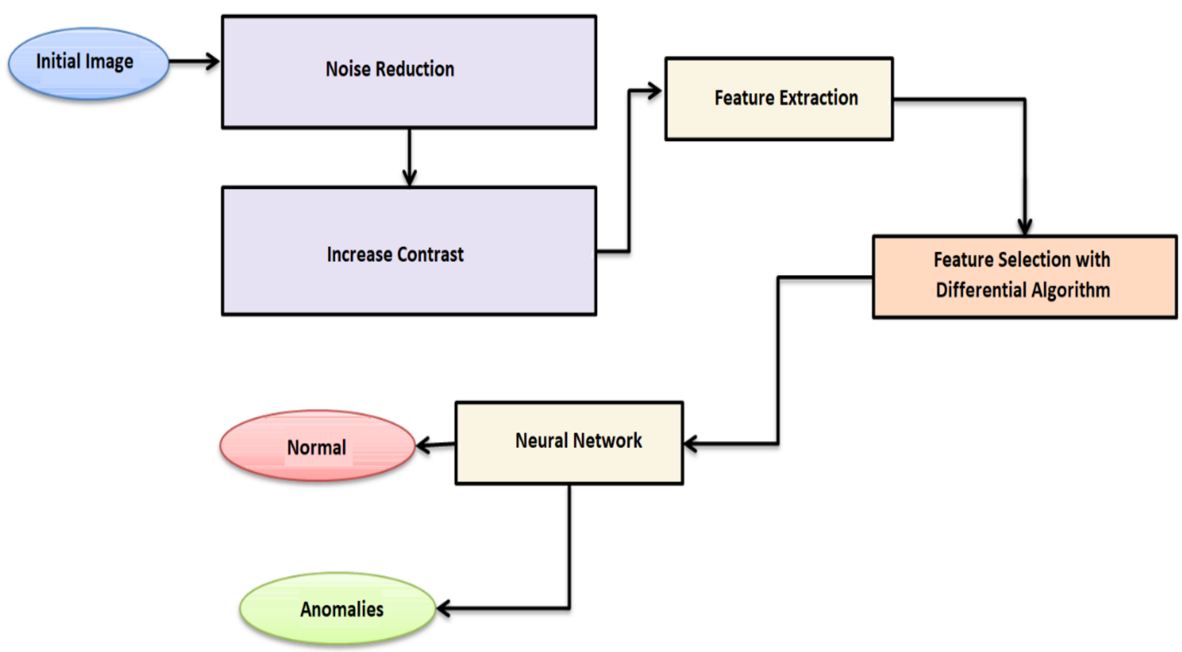

Fig. 2. Steps of the Proposed Method.

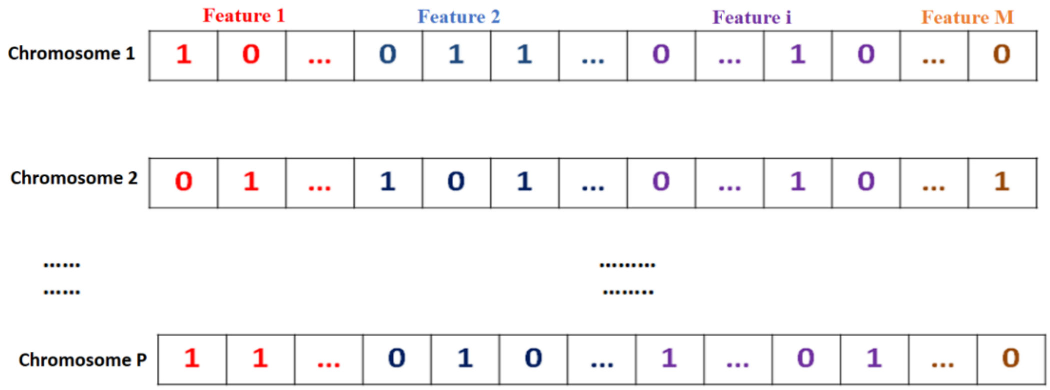

Fig. 3. Structure of Chromosomes in the Proposed Approach.

F. Feature Selection

Feature extraction is a process in which its salient and defining features are determined by performing an operation on the data. Feature extraction aims to make the raw data more usable for further statistical processing, classification, and clustering. Due to high convergence, the differential evolution algorithm was used at this stage to select basic features from among the extracted features. The dimensions of the images are 480 x 640. Due to the high and variable number of features in the images, this step on the image is necessary thing in order to be able to recognize the abnormality from the extracted features. In the differential evolution algorithm, a measure called mutual information according to Eq. (2) is used to evaluate (fitness calculation) of each chromosome produced in the algorithm:

$$MI(x;y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)\log\left(\frac{p(x,y)}{p(x)p(y)}\right) \quad (2)$$

where, generally, p(t) is the probability of variable t.

In order to further investigate and evaluate the performance of the differential evolution algorithm in the feature selection stage, in addition to using the mutual information merit function, the Merit criterion according to Eq. (3) was also used as the merit function in the proposed method (main fitness function).

$$Merit = \frac{k\overline{R_{cf}}}{\sqrt{k+k(k-1)R_{ff}}} \quad (3)$$

The differential evolution algorithm uses the generated training set to calculate the fitness function for the generated chromosome. Therefore, the evaluation of the chromosome to be in the fitness function includes the following three steps:

- A differential evolution algorithm codes the generated template.
- The selection of one of the two main images was randomly selected from the database and used by the differential evolution algorithm. In the next step, the evaluation begins.
- The output image produced by CNN to evaluate the chromosome fit value.

After generating a suitable template for CNN, all the images in two different databases were given to the CNN algorithm for classification. Also, the structure of chromosomes in the M-dimensional space of features is shown in Fig. 3. The features are coded as chromosomes in Fig. 3 to search in the space of the differential evolution algorithm. A unique code is considered for each feature in all P chromosomes.

G. Neural Network

In this research, the classification and diagnosis of the type of abnormality are based on the cellular neural network. For image processing purposes, the most common architecture is a two-dimensional network in which each processing unit interacts only with the M × N pixel cells that are its neighbors in the cell's permeation domain. Having a cellular network is a set of cells that has the property expressed in relation (4) C_{ij} for the cell $r \geq 0$ with radius $S_{ij}(r)$ neighbors:

$$S_{ij}(r) = \{C_{kl}: max(|k-i|,|l-j|) \leq r ,\ 1 \leq k \leq M ,\ 1 \leq l \leq N\} \quad (4)$$

That the stimulus $u_{ij}(t)$ which is generally not visible, the input $x_{ij}(t)$ using the mode c_{ij} of each cell that shows a measurable and observable value, and the input $y_{ij}(t)$ is external that enters the cell, the output is M × N using the coupled M × N differential equation called bias. The neural network dynamics describe the evolution of each cell's state and its interaction with its neighbors. For image processing purposes, the most common architecture is a two-dimensional network in which each processing unit interacts only with the M × N pixel cells that are its neighbors in the cell's permeation domain. Having a cellular network is a set of cells that has the property expressed in relation (5) C_{ij} for the cell r≥0 with radius $S_{ij}(r)$ neighbors:

$$S_{ij}(r) = \{C_{kl}: max(|k-i|,|l-j|) \leq r ,\ 1 \leq k \leq M ,\ 1 \leq l \leq N\} \quad (5)$$

The infiltration area with radius r = 1 corresponds to the neighborhood of 3x3, r = 2 to the neighborhood of 5x5 (Fig. 4).

That the stimulus $u_{ij}(t)$ which is generally not observable, the input $x_{ij}(t)$ using 4-time variables of the state c_{ij} of each cell that shows a measurable and observable value, and the input $y_{ij}(t)$ is external to the cell is input, the output is M × N using the coupled M × N differential equation called bias. The neural network dynamics $z_{ij}(t)$ is described as M × N redundancy, which describes the evolution of the state of each cell and its interaction with its neighbors. It is used to process an image with pixels. In order to classify humans automatically in M × N pixel images, the intensity value of each pixel is given to humans from a cellular neural network containing the suspicious area as the input of the corresponding cell in the cellular neural network. The neural network used in this method is a network with two neural layers and three neural layers, which has the function of activating linear hidden layers and sigmoid output. Table II shows the parameters used for the neural network algorithm.

The neural network used in this research is considered in Fig. 5. For this model, relation (6) holds true:

$$\hat{y}_j(t+1) = \sum_{i=1}^{l} G_i w_{ij} = \sum_{i=1}^{l} w_{ij} \exp(-\frac{\|x-m_i\|^2}{2\sigma_i^2}) \tag{6}$$

In this regard, $x(t) = [x_1(t) \dots x_m(t)]^T$ indicates the input vector, $\hat{y}_j(t+1)$, indicates the jth output, w_{ij}, indicates the synaptic weight between neurons i-th hidden and j-th output neuron, G_i, the Gaussian function in the i-th neuron of the hidden layer, m_i and σ_i^2 respectively represent the center and width of the Gaussian function and l, the number of the Gaussian function, or in other words, equal to the number of hidden layer nodes.

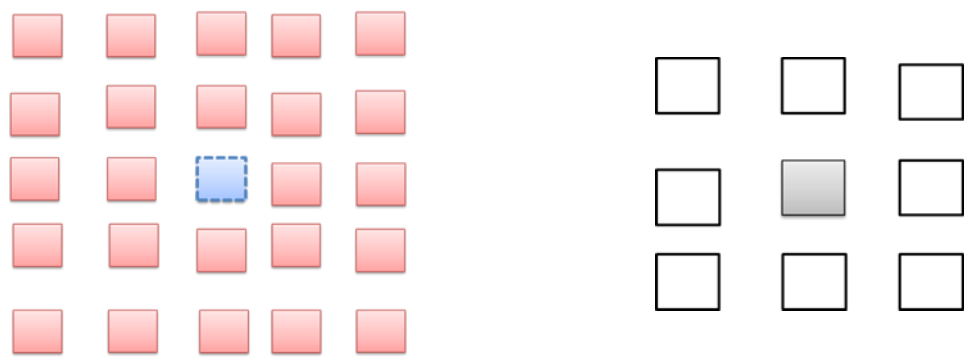

Fig. 4. Neighborhood of Cell c(i,j) r = 1, r = 2 Respectively.

TABLE II. PARAMETERS USED FOR NEURAL NETWORK ALGORITHM

arrangement	**Variable name**
Two layers and three layers	Network type
Variable	Number of hidden layers
Variable	the neurons
Linear and sigmoid	Stimulator function

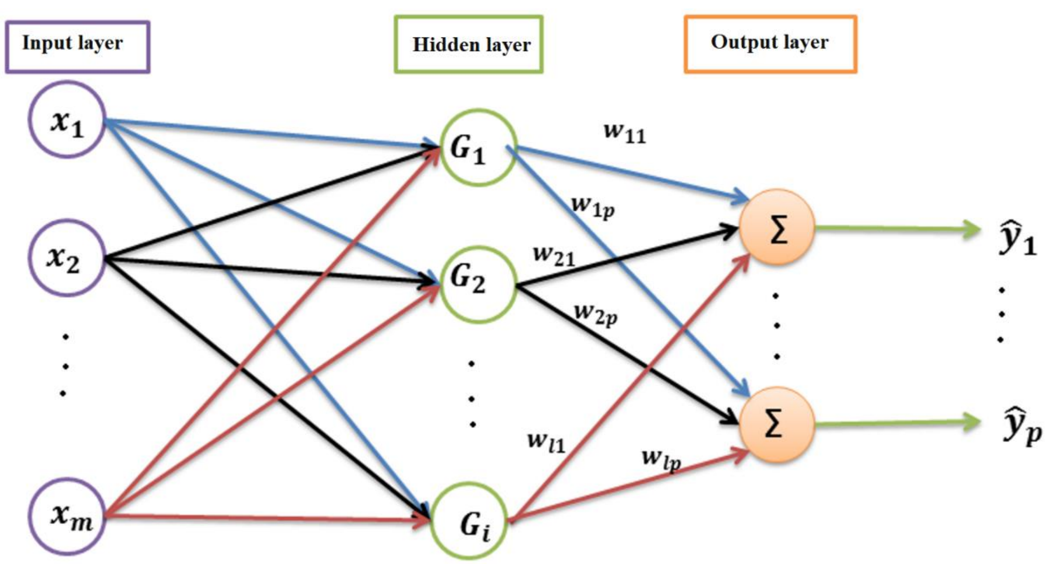

Fig. 5. Proposed Neural Network Architecture.

IV. Evaluation and Efficiency

In this section, the numerical results of the proposed methods are presented in comparison with similar algorithms. This comparison has been made for samples in different conditions and challenges. All the experiments of this research have been done with MATLAB version 2017b software in Windows 10 operating system and on a Dell laptop with a 2-core 2 GHz central processor and 2 GB internal memory.

A. Numerical Results Obtained

This research will analyze and select images from the UCSD Anomaly Detection Dataset database. Also, the presented method will be implemented using MATLAB software, and the results presented in this thesis will be compared with those presented in previous papers. In some cases, where there are both color and abnormal size lesions, thresholding can play an important role because it has a uniform illumination intensity. In contrast, the anomaly does not have uniform illumination dimensions. As a result, thresholding determines the approximate location of the abnormality better. It should be noted that small objects and objects that exist after thresholding can be removed by calculating their area. Fig. 6 shows the implementation of the proposed method on different images from the video.

The comparison of anomaly detection in training data as seen in Fig. 7 is shown for the accuracy of detection at different levels. And as the accuracy of diagnosis has increased, the amount of training data has taken a downward trend. The more training data there are, the better the model created using the suggested algorithm covers the images, resulting in a similar system with the arrival of new test data and the ability to detect anomalies. This is the reason for improving the accuracy of detection by increasing the number of training data.

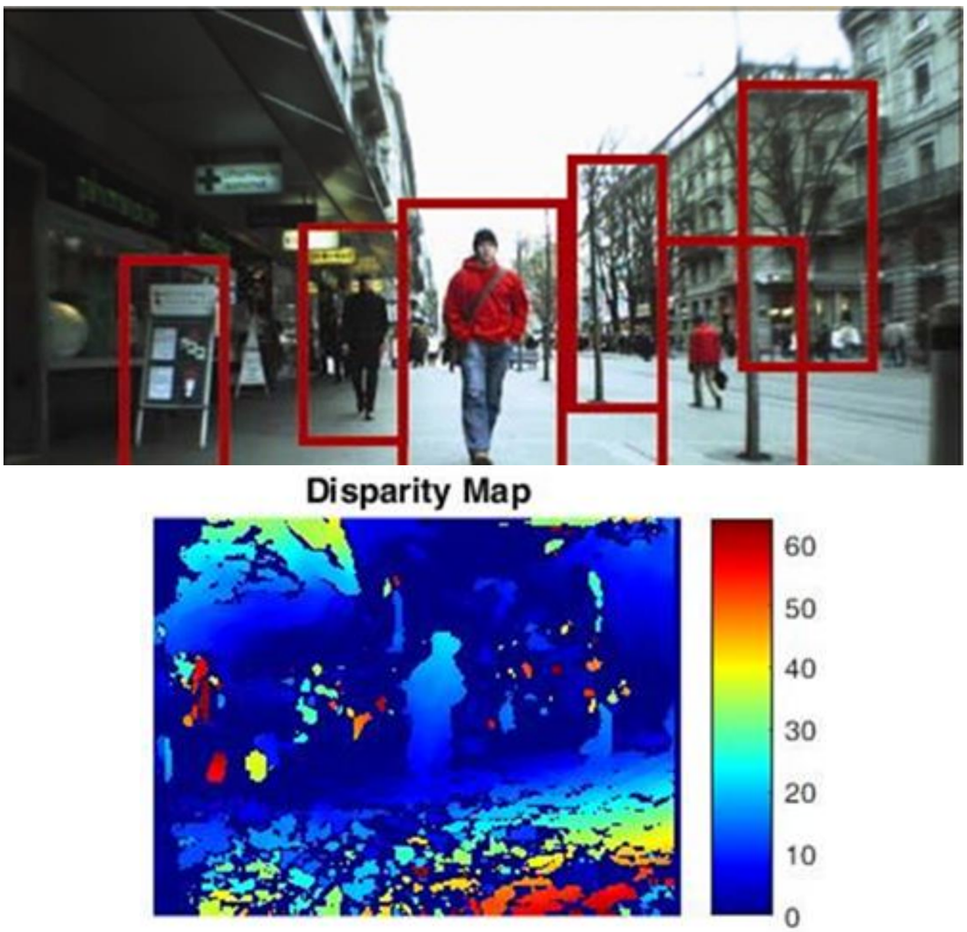

Fig. 6. Anomaly Detection using the Method in different Images from the Video.

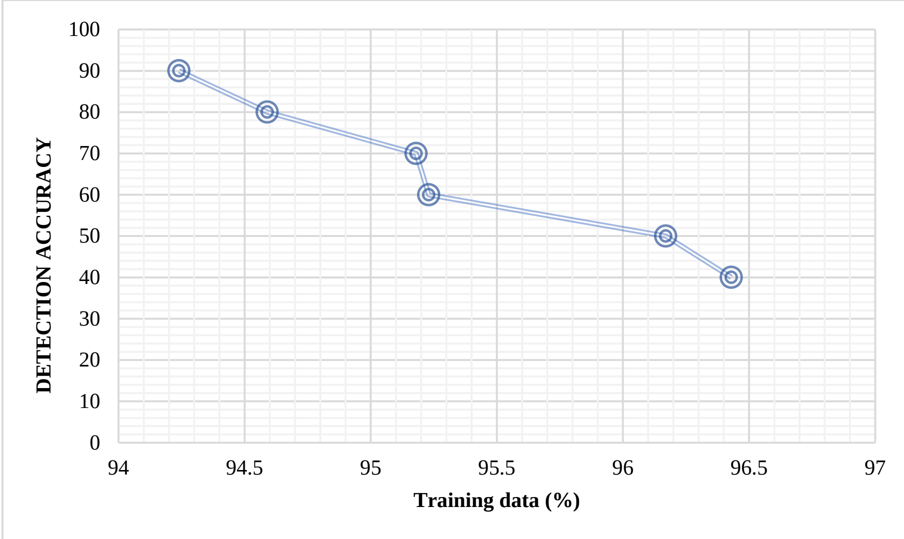

Fig. 7. Comparison of Anomaly Detection on Training Data.

TABLE III. ALGORITHM PARAMETERS

parameter	value
Initial population size	50
number of iterations (generations)	400
The number of features of each variable	20
Cutting rate	0/9
Composition rate	0/7
beta	0/08

Table III shows the parameters of the proposed difference evolution algorithm and their values. It should be noted that the values of the operators were selected by trial and error.

To ensure the correctness of the abnormality diagnosis, the final results and images must be reviewed and approved by an expert. However, to evaluate and check the efficiency of the proposed algorithm, the following criteria such as sensitivity, accuracy, and precision can be checked to estimate the correctness percentage of the proposed methods.

B. Comparison of Anomaly Estimates

Table IV shows the results of the proposed method. In our experiments, 40% to 90% of the training data were used, and we evaluated the anomaly detection results of the training data.

TABLE IV. ANOMALY ESTIMATION RESULTS WITH DIFFERENT NUMBER OF TRAINING DATA

Educational data (%)	Our Method
40	96.43
50	96.17
60	95.23
70	95.18
80	94.59
90	94.24

The decrease in detection accuracy with the increase in training data is because the more training data is, the more the model built with the proposed algorithm will have a higher computational load and perform anomaly detection with less accuracy. However, as it can be seen, the presented method is also effective in a small amount of training data. It is concluded from this figure that with fewer input data, we will have better training and better results will be obtained, and the prediction error will be minimized. It can be seen that even though all the parameters of this network are the same and only their initial weights and biases are different, the number of periods required to reach the desired error limit is very different. The interesting point is that, although after training, all the networks have reached the desired error limit and stopped at that limit, they will give different answers compared to the test samples.

C. ROC Curve

It has been used to evaluate and compare the tests presented in order to detect abnormalities with other methods in the same conditions. One of the ways to detect the strength of a classifier is to use the ROC chart and the area under that chart in such a way that if the area under the curve is low, it means a weak classification, and the more it is, it means a strong classification.

The difference between the ANN method and PROPOSED is in the combination of evolutionary algorithm and ANN, which parameters were optimized with differential evolution algorithm before neural network training. In Fig. 8, three classifications are compared based on the ROC curve. As can be seen, the sub-level of the proposed method is higher compared to the other two methods, and this means that the learning method has a better performance in classifying and diagnosing people's abnormalities.

D. Comparison of the Results of Several Methods for Abnormality Diagnosis

Accuracy and time criteria have been used to evaluate the presented algorithms to classify anomalies with other methods in the same conditions. Fig. 9 and 10, as well as Table V, compare the evaluation results of several methods for abnormality detection.

As seen in Fig. 9 and 10, the proposed solution was able to perform the anomaly detection operation in much less time and with relatively higher accuracy, which shows that the solution is optimal. This level of efficiency and optimality has been made possible with the help of a combined approach and the use of the differential evolution algorithm and the cellular neural network, which has found the solution to detect abnormalities in much less time and with proper accuracy. In the following, in the forms of 11 to 13 percent, the sensitivity, accuracy, and amount of correct prediction of the solutions are compared. The reason for checking these parameters is to evaluate the solution in the correct diagnosis with relatively high sensitivity. In diagnosing abnormal human behavior using video images, the sensitivity and correctness of the predictions must be high because the lower this level, the more reliable it is. It will also come down to the solution.

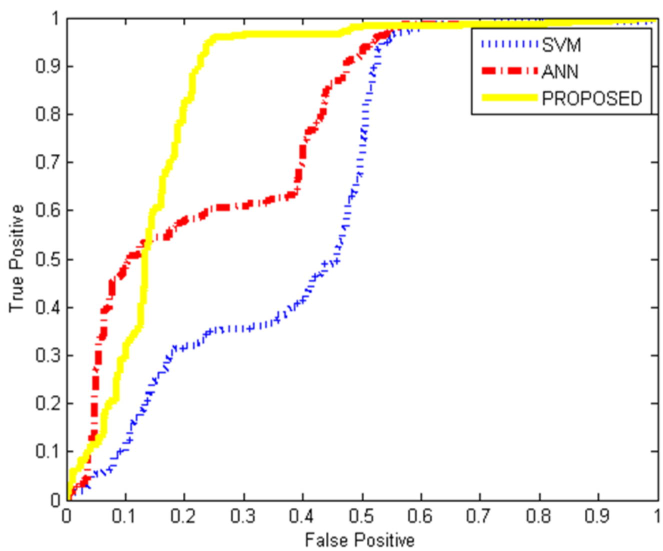

Fig. 8. Comparison of Methods.

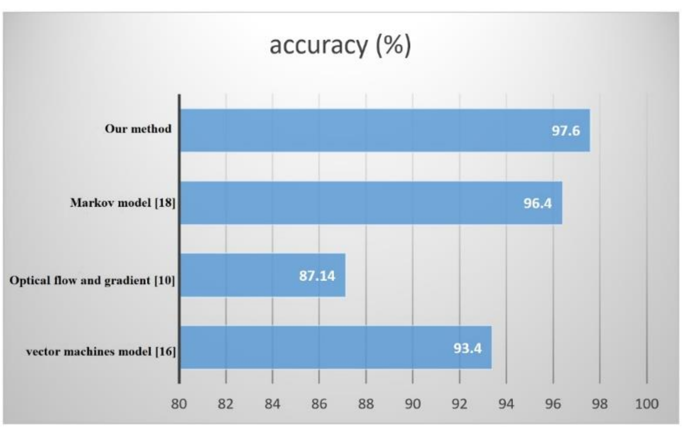

Fig. 9. Accuracy of Anomaly Detection in the Evaluated Solutions.

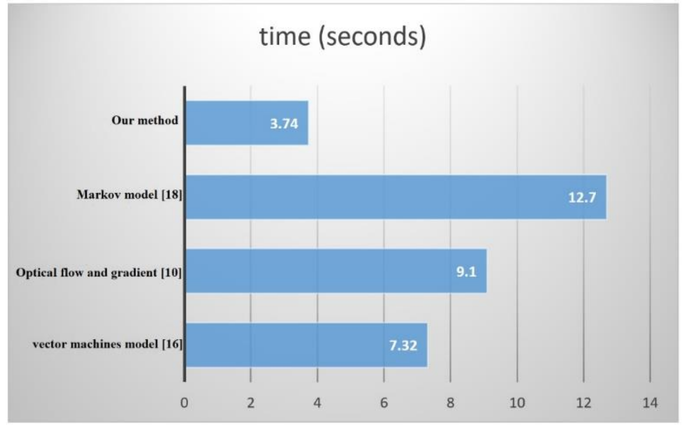

Fig. 10. Abnormality Detection Time in the Evaluated Solutions.

TABLE V. COMPARISON OF THE RESULTS OF SEVERAL METHODS FOR ABNORMALITY DIAGNOSIS

Method	accuracy (%)	time (s)
Vector machine model [16]	93.4	7.32
Optical flow and gradient [10]	87.14	9.1
Markov model [18]	96.4	12.7
Our method	**97.6**	3.74

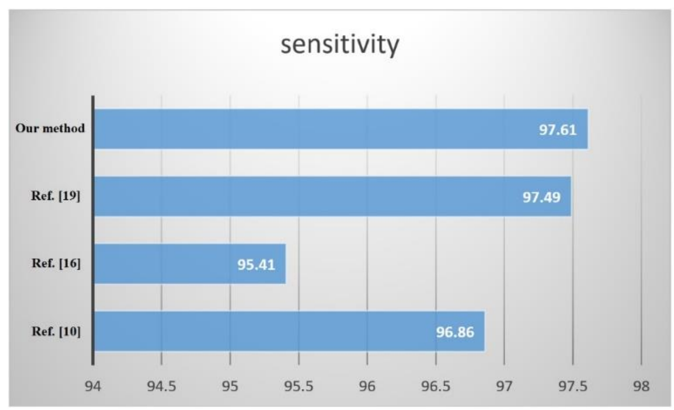

Fig. 11. The Level of Sensitivity in the Evaluated Solutions.

As seen in Fig. 11, the proposed solution has been able to perform better than all three solutions in terms of sensitivity. This level of optimization is possible due to the combined solution presented in the previous chapter.

In Fig. 12, the proposed solution has performed better than the two solutions presented in [19] and [10], but with a slight difference, it has performed weaker than the solution presented in [16], but considering that the method presented in this research In terms of other important parameters such as percentage of correct prediction or sensitivity, it has performed much better, this small weakness can be ignored.

Fig. 13 shows the percentage of correct predictions. As can be seen, the proposed solution was higher than the other evaluated solutions by a large margin. This level of optimality can compensate for the relative weakness of the solution in terms of accuracy. Table VI shows the numerical values of the proposed method in abnormality detection.

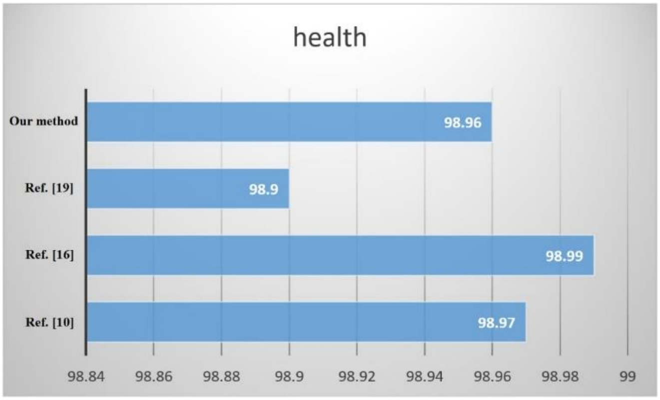

Fig. 12. The Degree of Accuracy in the Evaluated Solutions.

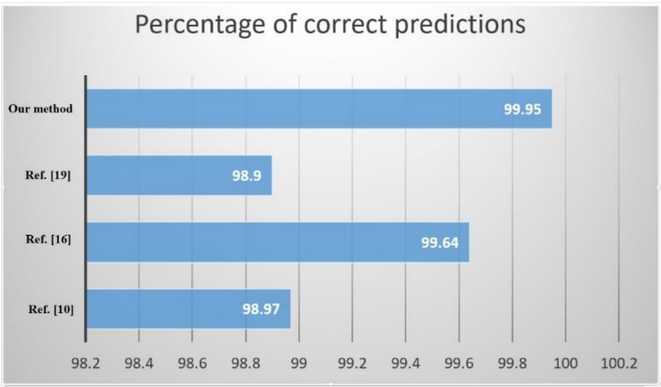

Fig. 13. Percentage of Correct Prediction in the Evaluated Solutions.

TABLE VI. NUMERICAL VALUES OF THE PROPOSED METHOD IN ABNORMALITY DETECTION

Ref.	health	sensitivity	Correct prediction
[10]	98/97	96/86	95/88
[16]	99/64	95/41	96/83
[19]	98/90	97/49	95/71
Our Method	98/96	97/61	96/54

The UCSD Anomaly Detection Dataset database was used to analyze the images.

http://www.svcl.ucsd.edu/projects/anomaly/dataset.html.

V. CONCLUSION

In this research, a computationally efficient algorithm for anomaly detection on video images with high resolution is presented. The proposed method has higher speed and accuracy compared to other methods. The advantage of the algorithm is that it has an execution time of three seconds on a home computer. The proposed method has higher accuracy and precision in comparison with other methods. The advantage of the algorithm is that it has an accuracy of 96.5% on a home computer, and the average sensitivity criterion is 98.6%/97.2%.

VI. SUGGESTIONS

It is suggested that in the future, the following things should be investigated to improve the methods presented in this paper. The proposed method has a good potential for classifying video objects such as cars in color images. Improving the proposed methods to reduce the time of abnormality detection, which can be used to improve the abnormality detection using the combined methods of the neural network, which requires a supercomputer.

REFERENCES

[1] Vats, E., & Chan, C. S. (2016). Early detection of human actions—A hybrid approach. Applied Soft Computing, 46, 953-966.

[2] Narang, V., & Solanki, A. (2023). An Efficient Algorithm for Human Abnormal Behaviour Detection Using Object Detection and Pose Estimation. In The Fifth International Conference on Safety and Security with IoT (pp. 47-64). Springer, Cham.

[3] Y. Cong, J. Yuan, and Y. Tang, "Video anomaly search in crowded scenes via spatio-temporal motion context," IEEE Trans. Inf. forensics Secur., vol. 8, no. 10, pp. 1590–1599, 2013.

[4] Guha, A., & Samanta, D. (2021). Hybrid approach to document anomaly detection: an application to facilitate RPA in title insurance. International Journal of Automation and Computing, 18(1), 55-72.

[5] V. Kaltsa, A. Briassouli, I. Kompatsiaris, L. J. Hadjileontiadis, M. G. Strintzis, Swarm intelligence for detecting interesting events in crowded environments, IEEE Trans.ImageProcess.24(7)(2015)2153–2166.

[6] Y. Shenn, W. Hu, M. Yang, J. Liu, and C. T. Chou, "Efficient background subtraction for tracking in embedded camera networks," in Information Processing in Sensor Networks (IPSN), 2012 ACM/IEEE 11th International Conference on, 2016, pp. 103–104.

[7] J. M. Chaquet, E. J. Carmona, and A. Fernández-Caballero, "A survey of video datasets for human action and activity recognition," Comput. Vis. Image Underst., vol. 117, no. 6, pp. 633–659, 2013.

[8] Sun, J., Zhang, Y., & Trik, M. (2022). PBPHS: A Profile-Based Predictive Handover Strategy for 5G Networks. Cybernetics and Systems, 1-22.

[9] Chapnevis, A., Güvenç, I., & Bulut, E. (2020, November). Traffic Shifting based Resource Optimization in Aggregated IoT Communication. In 2020 IEEE 45th Conference on Local Computer Networks (LCN) (pp. 233-243). IEEE.

[10] A. Oliver, J. Freixenet, J. Marti, E. Pérez, J. Pont, E. R. E. Denton, and R. Zwiggelaar, "A review of automatic mass detection and segmentation in mammographic images," Med. Image Anal., vol. 14, no. 2, pp. 87–110, 2017.

[11] Ali, M. M., Noorain, S., & Qaseem, M. S. (2023). Suspicious Human Behaviour Detection Focusing on Campus Sites. In Emerging IT/ICT and AI Technologies Affecting Society (pp. 171-183). Springer, Singapore.

[12] Rafiee, P., & Mirjalily, G. (2020). Distributed Network Coding-Aware Routing Protocol Incorporating Fuzzy-Logic-BasedForwarders in Wireless Ad hoc Networks. Journal of Network and Systems Management, 28(4), 1279-1315.

[13] Y. Xu, X. Cao, and H. Qiao, "An efficient tree classifier ensemble-based approach for pedestrian detection," Syst. Man, Cybern. Part B Cybern. IEEE Trans., vol. 41, no. 1, pp. 107–117, 2011.

[14] Sundaramoorthy, S. U. R. I. Y. A. (2021). Deep Learning Based Hybrid Approach For Facial Emotion Detection.

[15] Trik, M., Mozaffari, S. P., & Bidgoli, A. M. (2021). Providing an adaptive routing along with a hybrid selection strategy to increase efficiency in NoC-based neuromorphic systems. Computational Intelligence and Neuroscience, 2021.

[16] H. Mousavi, S. Mohammadi, A. Perina, R. Chellali, V. Murino, Analyzing tracklets for the detection of abnormal crowd behavior,in:IEEE Winter Conference Applications of Computer Vision(WACV) ,IEEE, Hawaii, 2015,pp. 148–155.

[17] S. Zhang, R. Benenson, and B. Schiele, "Filtered channel features for pedestrian detection," in Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on, 2015, pp. 1751–1760.

[18] Lotfi, F., Semiari, O., & Saad, W. (2021). Semantic-Aware Collaborative Deep Reinforcement Learning Over Wireless Cellular Networks. arXiv preprint arXiv:2111.12064.

[19] Hashemzehi, R., Mahdavi, S. J. S., Kheirabadi, M., & Kamel, S. R. (2020). Detection of brain tumors from MRI images base on deep learning using hybrid model CNN and NADE. biocybernetics and biomedical engineering, 40(3), 1225-1232.

[20] A. Wiliem, V. Madasu, W. Boles, and P. Yarlagadda, "A suspicious behaviour detection using a context space model for smart surveillance systems," Comput. Vis. Image Underst., vol. 116, no. 2, pp. 194–209, 2012.

[21] Chenarlogh, V. A., Razzazi, F., & Mohammadyahya, N. (2019, December). A multi-view human action recognition system in limited data case using multi-stream CNN. In 2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS) (pp. 1-11). IEEE.

[22] Trik, M., Molk, A. M. N. G., Ghasemi, F., & Pouryeganeh, P. (2022). A Hybrid Selection Strategy Based on Traffic Analysis for Improving Performance in Networks on Chip. Journal of Sensors, 2022.

[23] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 1, pp. 18–32, 2014.

[24] Golrou, A., Sheikhani, A., Nasrabadi, A. M., & Saebipour, M. R. (2018). Enhancement of sleep quality and stability using acoustic stimulation during slow wave sleep. International Clinical Neuroscience Journal, 5(4), 126.

[25] M. J. Roshtkhari and M. D. Levine, "An on-line, real-time learning method for detecting anomalies in videos using spatio-temporal compositions," Comput. Vis. Image Underst., vol. 117, no. 10, pp. 1436–1452, 2013.

[26] O. P. Popoola and K. Wang, "Video-based abnormal human behavior recognition—a review," IEEE Trans. Syst. Man, Cybern. Part C (Applications Rev., vol. 42, no. 6, pp. 865–878, 2014.

[27] Trik, M., Pour Mozafari, S., & Bidgoli, A. M. (2021). An adaptive routing strategy to reduce energy consumption in network on chip. Journal of Advances in Computer Research, 12(3), 13-26.

[28] H. Nallaivarothayan, D. Ryan, S. Denman, S. Sridharan, and C. Fookes, "An evaluation of different features and learning models for anomalous event detection," in Digital Image Computing: Techniques and Applications (DICTA), 2013 International Conference on, 2013, pp. 1–8.

[29] X. Zhu, J. Liu, J. Wang, C. Li, and H. Lu, "Sparse representation for robust abnormality detection in crowded scenes," Pattern Recognit., vol. 47, no. 5, pp. 1791–1799, 2014.

[30] Kunkel, G., Madani, M., White, S. J., Verardi, P. H., & Tarakanova, A. (2021). Modeling coronavirus spike protein dynamics: implications for immunogenicity and immune escape. Biophysical Journal, 120(24), 5592-5618.

[31] Shifu Zhou, Wei. Shen, Dan. Zeng, Mei. Fang, Yuanwang. Wei,Zhijiang. Zhang, Spatial–temporal convolutional neural networks for anomaly detection and localization in crowded scenes, Signal Processing: Image Communication 47 (2016), pp. 358–368.

[32] X. Liu, G. Zhao, J. Yao, and C. Qi, "Background subtraction based on low-rank and structured sparse decomposition," Image Process. IEEE Trans., vol. 24, no. 8, pp. 2502–2514, 2015.

[33] Hosseini, S. (2022). The role of pertomix approaches in early detection of cancer. Proteomics, 9(01), 177-186.

[34] A. A. Sodemann, M. P. Ross, and B. J. Borghetti, "A review of anomaly detection in automated surveillance," IEEE Trans. Syst. Man, Cybern. Part C (Applications Rev., vol. 42, no. 6, pp. 1257–1272, 2016.

[35] Rezaee, K., Rezakhani, S. M., Khosravi, M. R., & Moghimi, M. K. (2021). A survey on deep learning-based real-time crowd anomaly detection for secure distributed video surveillance. Personal and Ubiquitous Computing, 1-17.

[36] Charandabi, S., & Ghanadiof, O. (2022). Evaluation of Online Markets Considering Trust and Resilience: A Framework for Predicting Customer Behavior in E-Commerce. Journal of Business and Management Studies, 4(1), 23-33.

[37] M. Errami and M. Rziza, "Improving Pedestrian Detection Using Support Vector Regression," in 2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV), 2016, pp. 156–160.

[38] Usman, I., & Albesher, A. A. (2021). Abnormal Crowd Behavior Detection Using Heuristic Search and Motion Awareness. International Journal of Computer Science & Network Security, 21(4), 131-139.

[39] Sabha, A., & Selwal, A. (2021, September). HAVS: Human action-based video summarization, Taxonomy, Challenges, and Future Perspectives. In 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1-9). IEEE.

[40] Mu, H., Sun, R., Yuan, G., & Wang, Y. (2021). Abnormal Human Behavior Detection in Videos: A Review. Information Technology and Control, 50(3), 522-545.

[41] Radhoush, S., Shabaninia, F., & Lin, J. (2018, February). Distribution system state estimation with measurement data using different compression methods. In 2018 IEEE Texas Power and Energy Conference (TPEC) (pp. 1-6). IEEE.

[42] Gill, R., Singh, J., & Modgill, A. (2022, March). A Retrospective CNN-LSVM Hybrid Approach for Multimodal Emotion Recognition. In 2022 International Conference on Decision Aid Sciences and Applications (DASA) (pp. 1281-1285). IEEE.

[43] HassanVandi, B., Kurdi, R., & Trik, M. (2021). Applying a modified triple modular redundancy mechanism to enhance the reliability in software-defined network. International Journal of Electrical and Computer Sciences (IJECS), 3(1), 10-16.

[44] Ramprasadi, N., Shah, P., & Vyas, D. (2020, May). Hybrid Approach For Real Time Crowd Activity Identification Using Segmentation. In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 390-394). IEEE.

[45] Trik, Mohammad, Amir Massoud Bidgoli, Hossein Vashani, and Saadat Pour Mozaffari. "A new adaptive selection strategy for reducing latency in networks on chip." Integration (2022).

[46] Vahidi Farashah, M., Etebarian, A., Azmi, R., & Ebrahimzadeh Dastjerdi, R. (2021). A hybrid recommender system based-on link prediction for movie baskets analysis. Journal of Big Data, 8(1), 1-24.