Abnormal Crowd Behavior Detection Using Heuristic Search and Motion Awareness

Imran Usman[1†] and Abdulaziz A. Albesher[2††],

College of Computing and Informatics, Saudi Electronic University, Riyadh, Saudi Arabia

Summary
In current time, anomaly detection is the primary concern of the administrative authorities. Suspicious activity identification is shifting from a human operator to a machine-assisted monitoring in order to assist the human operator and react to an unexpected incident quickly. These automatic surveillance systems face many challenges due to the intrinsic complex characteristics of video sequences and foreground human motion patterns. In this paper, we propose a novel approach to detect anomalous human activity using a hybrid approach of statistical model and Genetic Programming. The feature-set of local motion patterns is generated by a statistical model from the video data in an unsupervised way. This features set is inserted to an enhanced Genetic Programming based classifier to classify normal and abnormal patterns. The experiments are performed using publicly available benchmark datasets under different real-life scenarios. Results show that the proposed methodology is capable to detect and locate the anomalous activity in the real time. The accuracy of the proposed scheme exceeds those of the existing state of the art in term of anomalous activity detection.

Key words:
abnormal behavior detection, genetic programming, crowd analysis, motion pattern, anomaly

1. Introduction

Consider a smart scenario in which a robot is vigilant in identifying all irregular human behavior, preventing crowd turbulence and violence until they worsen or intensify, serving as a monitoring agent in public and private spaces, preventing robberies/theft by notifying concerned officials. Though we may not living in such era, prodigies are working hard to developed such intelligent machines which lead us towards that stage where such sophisticated robots work rapidity without human interference.

The increase in threats to security and safety is one of the most demanding problems in the world and needs better security plans and executions using sophisticated techniques to avoid un-pleasant incidents [1]. Surveillance systems are one such tool that is commonly used to observe several locations on a computer and alert the appropriate authorities if any suspicious or anomalous behavior is detected [2]. Wide screens with a variety of video sequences corresponding to the number of cameras deployed are needed for such monitoring. Owing to human operator limits, it is not possible to constantly track any of the videos on the computer at all times. As a result, several actionable events go undetected, and the video archive is only used for evidence purpose.

For suspicious activity detection in modern security systems, there is a change from a human operator model to a machine-assisted intelligent surveillance model in order to assist human operators and react to an unexpected incident in real time. This is performed by automatically detecting the anomalies in the videos from different cameras. It is the challenging task for the machine due to the vast amount of data to be analyzed from various cameras and the random movements in each frame. The visual contents analysis includes scene understanding and distinguishing the normal and abnormal activities in the scene [3]. There are numerous vital limitations in detection of anomaly using video surveillance platforms. One of the limitations is that, there is no basic accepted definition of anomaly because it varies from scene to scene [4]. For example, running the shopping mall would be considering abnormal behavior while running on the zebra crossing would be normal. Such limitation made computer based approaches difficult to identify anomalies in the video frames. To automatically detect anomalies, there is a need for an efficient and fast algorithm to imply in real time videos.

Since crowd activities and its understanding are difficult to model, automatic surveillance in busy spaces is a challenging task. For this, the existing literatures classified in to two approaches are, namely, (i) the object-based approach and (ii) the holistic approach. In the object-based approach, the crowd is treated as a set of individual entities or objects. Each entity is tracked for its motion and its subsequent analysis. However, identifying objects, observing trajectories, and understanding actions in crowded scenes are difficult for these object-based schemes [5]. The holistic approach, on the other hand, treats the crowd as a single entity and monitors its movement and action/behavioral patterns. These approaches use common features like spatial–temporal gradients or optical flow data to locate regions of dramatic motion. Anomaly identification is carried out using pre-trained classifiers by modeling normal/abnormal crowd motion patterns.

Manuscript received April 5, 2021
Manuscript revised April 20, 2021
https://doi.org/10.22937/IJCSNS.2021.21.4.18

The work is based on holistic approach. In this paper, we propose a scheme to efficiently analyze the video, extract different activities, interactions and classify normal and abnormal behaviors. It is performed with the help of statistical modeling of the video and advanced machine learning approach. In this paper, genetic programming (GP) based classifier has been proposed to classify different behaviors efficiently. To our best knowledge, it is for the first time to use genetic programming in crowd analysis for classifying normal and abnormal behaviors.

The rest of the paper is organized as follows: section 2 includes the existing literature. Section 3 presents the proposed methodology with the GP based modeling. Section 4 consists of experimental results and discussion and conclusion is given in the section 5.

2. Literature Review

In terms of current learning models, anomaly detecting is treated as a binary classification problem, with activities categorized as normal or abnormal. The area of anomaly detection in videos has a considerable amount of literature, and scientists have conducted important survey papers on the topic [1-4]. Researchers have proposed different methodologies according to different events. Aguilar et al. [5], proposed inter-frame-pixel motion approach to detect the anomalous crowd behavior. In [6], crowd's energy measurement based approach is used to predict the anomalous behavior. Image's pixels are considered as particles. Then, the energy of these particles is measured by optical flow method. Ravanbakhsh et al. [7], introduced Artificial neural network architecture to efficiently extract key features change across time. The temporal change in the CNN features is considered the clue for anomaly in the crowd.

The high density of crowd in the scene makes system miss the key individual under consideration. A dual artificial neural network architecture is presented in [8] for anomaly detection in the dense crowd. The dual LSTM network is trained using key features obtained by density-heat-map and optical flow algorithms. Chaudary et al.[9] proposed a framework which can identify various anomalous crowd behavior in real time visual contents. In [9], GMM (Gaussian Mixture Model) is used for object detection and rule based approach is used for anomaly detection. Sultani et al. [10] proposed a neural network for abnormal activity identification by using normal and abnormal video as bags which is fed to the neural network for training. The highest ranking lost function defines the highest scores for the anomalous instance.

In [11], a hybrid approach of multiclass SVM and 3DCNN is proposed where the spatio-temporal features are extracted using neural network and anomalous scenes are classified using multiclass SVM classifier. Most of the classification tasks are accomplished using transfer learning. In transfer learning, pre-trained neural network is used to avoid extensive machine training and huge input data. In [12,13] , pre-trained fully connected CNN algorithms are proposed for global anomaly identification in a scene.

Navneet et al. [14] proposed a methodology based on correlation analysis of the optical flow for accurate and fast detection of anomalous behavior of a crowd. Further, several factors such as frame gap (the gap of frames for processing during event detection) and illumination condition were also studied by [14]. The correlation coefficient of the optical flow of consecutive frames provides a pattern of abnormal and normal activity. Afiq et al. [15] conduct a comprehensive study of analyzing the crowd features such as density, motion and trajectory to detect any abnormality in the crowd. They also aimed to provide details of several detection algorithms including Spatio-Temporal Technique (STT), Hidden Markov Model (HMM),Optical Flow (OF) method and Gaussian Mixture Model (GMM).

In Most of the studies, the abnormal event detection problem is based on outlier detection strategy where a model is trained using familiar events. Radu et al. [16] proposed a method of abnormal event detection as a multi-class classification task. A one-versus-rest abnormal event classifier is applied to distinguish each normality cluster from the rest. Jun et al. [17] proposed a new method for abnormal behavior detection using deep learning. Two State De-noising Auto Encoders (SDAEs) is utilized to automatically learn appearance feature and motion feature respectively, which are constrained in the space–time volume along dense trajectories that carry rich motion information to reduce the computational complexity. A Hybrid frame work (OF-ConvAE-LSTM) is proposed in [18] to detect anomalous crowd behavior.

The (OF-ConvAE-LSTM) algorithm detects anomalies using Convolutional Autoencoder and Convolutional Long Short-Term Memory in an unsupervised manner. In order to extract velocity and direction information features of foreground objects to train the model, a dense optical flow model is applied. In [19], an ANN model is proposed for detection of abnormal crowd behavior by amplifying the variances between projecting and real frames in moving object regions in videos. The study of Karishma et al. [20] includes analyzing and summarization of deep learning techniques for video based anomaly detection. They presented the comparison of ANN algorithms in anomaly detection process using video from both the view-point of accuracy-oriented approaches and real-time processing oriented approaches. Their findings demonstrated that the use of deep learning for anomaly detection has achieved remarkable results on both the accuracy oriented and real time processing oriented objectives of anomaly detection. In [21], a Neural Network model is used to detect anomalous crowd behavior. The

CNN is trained using hand crafted features of Motion information images which are calculated using optical flow method. Sing et al. [22] proposed transfer learning scheme where numbers of pre-trained neural network models are fine tuned to generate best features vectors which are next fed to binary classifiers to classify normal and abnormal crowd behavior. In [23], a pre-trained VGG16 ANN is used along with the LSTM model to predict the abnormal activity features. The scheme requires huge computational power to learn required anomalous behavior features. This high processing limitation makes the system unable to be applied on real time video surveillance. To address the challenges of detecting abnormal activity in real time, spatiotemporal auto encoder techniques have been proposed in [24, 25].

The meaning of abnormal behavior cannot be limited to viciousness or other rare events. There are some services like daycares and crèches which have their own set of abnormal activities. To address the child abuse issue of daycares and crèches and to detect the abnormal activities in these services, a systematic study is proposed in [26]. A lot of literature has been done on abnormal activity detection and promising results have been produced in these studies. But the existing works are limited to specific abnormal event detection because the proposed schemes in the studies are utilized specific event data. Abnormal activity detection is a challenging task because there are various natures of abnormalities. A large dataset of various abnormal events and abnormal events evaluation criteria are introduced in [27].

In the proposed framework, which is an enhancement of the work presented in [31], the motion of the object is determined using gradient based approach in spatio-temporal volumes (cuboids). The mean and variance of the objects motion are measured in the adjacent cuboids. In addition, we employ the idea of motion aware feature by utilizing temporal augmented network using a cascade of optical flow maps as presented in [32]. This features map is given as input to the GP based enhanced classifier with sigmoid activation to classify normal and abnormal behavior of the crowd in the scene. The proposed technique is efficient and outperforms contemporary works in literature in terms of classification accuracy and temporal cost.

3. Proposed Automated Crowd Behavior Detection Scheme

2.1 Abnormal Behavior: Some Challenges

Predicting abnormal behavior in crowded motion scenes is difficult due to the vast number of objects, lighting conditions, weather effects, erratic motion, occlusion, and scene limits including borders. The main aim of this work is to use accurate modeling and classification to classify various events in the scene. The concept of cuboids is used to discriminate numerous activities across the time varying video sequence as shown in figure 1. A video sequence is divided into small sequence of video clips. Each video clip is then further subdivided spatially to create spatio-temporal volumes containing various events. A gradient based approach [28], is then utilized to model the motion of each object.

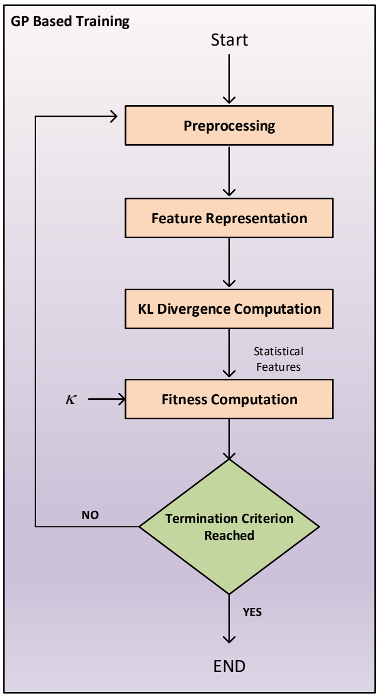

Fig. 1. General architecture of the GP based training phase

2.2 Statistical Modeling and Feature Representation

The irregular and non-uniform motion patterns are represented by gradient based method. Since in the proposed work cuboids concept is used, therefore, the gradient will operate in three dimensional (3D) space, horizontal, vertical and time represented by (x, y, t). For a particular pixel i in a cuboid I the 3D gradient is represented as:

$$\nabla I_i = \left[I_{i,x}, I_{i,y}, I_{i,t} \right] = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial t} \right]^T \quad (1)$$

Where x, y and t represent the horizontal, vertical and time dimensions of the cuboid. The gradient of each pixel within a cuboid defines the sequence of moving objects in a spatio-temporal volume. As we statistically model this gradient distribution, it conforms to a 3D Gaussian distribution $\mathrm{N}(\mu, \Sigma)$, represented by mean (μ) and covariance (Σ)

$$\mu = \frac{1}{N}\sum_{i}^{N} \nabla I_i \tag{2}$$

And a covariance matrix (Σ) is given by,

$$\Sigma = \frac{1}{N}\sum_{i}^{N} (\nabla I_i - \mu)(\nabla I_i - \mu)^T \tag{3}$$

The notation $O_t^{x,y}$ is used to represent the spatial and temporal motion characterized by $\mu_t^{x,y}$ and $\Sigma_t^{x,y}$ for a spatial location (x, y) and time t.

We use Kullback-Leibler (KL) divergence [29] to calculate the motion difference from one cuboid to its neighboring cuboid and the resemblance between the adjacent cuboids based on the statistical representation of a cuboid.

We can find identical cuboids inside a scene using the KL divergence technique. The larger the resemblance between cuboids, the less divergence there is. As a result, we can conclude that their motion patterns are strikingly identical. The cuboids are then clustered together to create a cluster that represents a single activity class. Clustering algorithms in a traditional scenario imply that the number of clusters be known ahead of time. On the other hand, prototypical movements in motion sequences are impossible to predict and are heavily dependent on crowd density, object motion and direction, video clip scale, and other factors. As a result, to merge adjacent cuboids, an online clustering technique is used.

The next step is to calculate the symmetric KL divergence between a cuboid and each prototype P_s. If the divergence between the cuboid and prototypes exceeds a predetermined threshold for entirely prototypes $\{P_s \mid s = 1, 2...., S\}$, then the new cuboid $O_t^{x,y}$ is regarded as a unique prototype. The prototype P_s with the smallest divergence or a KL divergence less than the threshold is modified according to the following rule:

$$P_s = \frac{1}{N_s + 1} O_t^{x,y} + \left(1 - \frac{1}{N_s + 1}\right) P_s \tag{4}$$

In above equation, N_s represents the total number of cuboids in prototype P_s. The weighted sum of actual data in the cuboids is reflected in equation 4 since the cuboids and prototypes are multivariate Gaussian distributions. As a result, the expectation centroid definition, as provided in [30], is used, where mean and variance defined as:

$$\mu_c = \frac{1}{N}\sum_{n=1}^{N} E[x_n] = \frac{1}{N}\sum_{n=1}^{N} \mu_n \tag{5}$$

and,

$$\begin{aligned}\Sigma_c &= \frac{1}{N}\sum_{n=1}^{N} E\left[(x_n - \mu_n)(x_n - \mu_n)^T\right] \\ &= \frac{1}{N}\sum_{n=1}^{N}\left[\Sigma_n + \mu_n\mu_n^T - \mu_c\mu_c^T\right]\end{aligned} \tag{6}$$

Furthermore, the KL divergence of each cuboid is used to determine how similar different cuboids are. To account for the results of neighboring motions, we take the mean and variance of KL divergences for the N neighbor cuboids. Such pattern characterizes the steady-state motion. Kratz and Nishino [28] used coupled HMMs to capture neighboring movements based on the cuboid. We use basic mean and variance calculated in the temporal and spatial directions to deal with the complexity of this method. The mean of a $N \times N$ filter is calculated as follows:

$$\mu_{x,y} = \frac{\sum_{s=-a}^{a}\sum_{t=-b}^{b} d_{KL}(x+a, y+b)}{\sum_{s=-a}^{a}\sum_{t=-b}^{b}(1)} \tag{7}$$

And variance of each filter is,

$$\sigma_{x,y}^2 = \frac{\sum_{s=-b}^{b}\left(\sum_{t=-b}^{b}\left(d_{KL}(x+a, y+b) - \mu_{x,y}\right)^2\right)}{\sum_{s=-a}^{b}\sum_{t=-b}^{a}(1)} \tag{8}$$

where, d_{KL} is the KL divergence.

In the following subsection, we explain the GP based classifier and the use of enhanced feature-set and learning mechanism with activation.

2.3 Enhanced GP Classifier with Activation Function

In order to learn the intrinsic scene peculiarities and the dependencies between various actors and objects in the scene, we employ GP with an activation function. The implementation details of the GP are similar to those presented in [31] with the addition of an activation function during GP evolution and the idea of incorporating a motion aware feature in the statistical features as presented in [32].

The advantage of adding an activation function is to speed up the learning and produce a number of possible optimal solutions towards the class boundaries. Once a mathematical expression trained in this manner, the results are more generalized and yields better performance for real world applications. In addition to this, the use of a temporal augmented network to learn a compact motion-aware feature, fully exploits a number of temporal constraints that contributes significantly towards the detection of anomalies across the time axis.

For the training in GP, we need to define appropriate GP terminal set and the function set. The function set for the proposed technique includes SIN, COS, LOG, +, -, *, protected division, MAX and MIN functions. The terminal set forms the leaf nodes of the GP tree data structures and can be further divided into constant terminals and the variable terminals. We use random constants in the range of [-1, 0] to make up the constant terminals. The variable terminal set comprises the features generated in the previous subsections. Their selection is one of the core components in the evolution of the optimal solutions. We use the mean in (x, y) direction ($\mu_{x,y}$), the mean in t direction (μ_t), variance in (x, y) direction ($\sigma^2{}_{x,y}$), the variance in t direction ($\sigma^2{}_t$), and the motion aware feature matrix κ. Thus, the fitness score, ξ, is calculated as:

$$\xi = Eval(F(\kappa, \mu_{x,y}, \mu_t, \sigma^2{}_{x,y}, \sigma^2{}_t)) \tag{9}$$

Next, we apply the activation function in order to activate the result towards a binary decision. We have used the sigmoid activation function given by the following relation:

$$S(\xi) = \frac{1}{1+e^{-\xi}} \tag{10}$$

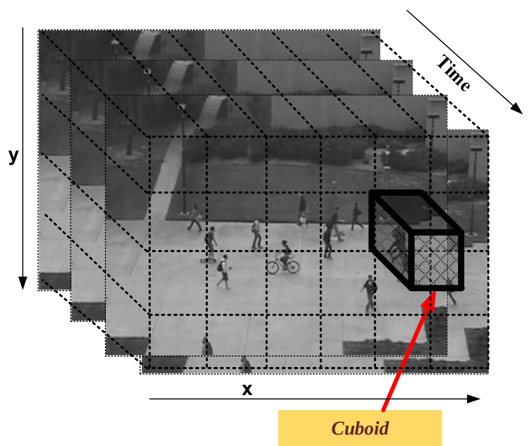

Fig. 2. Example of a spatio-temporal cuboid in a video scene from UCSDped2 dataset.

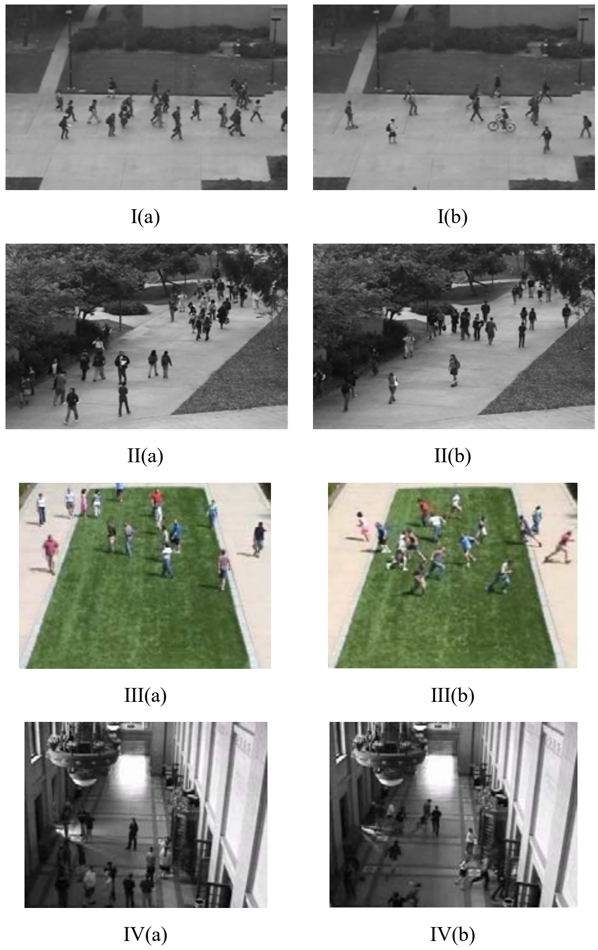

Fig. 3. Example video screen shots from [3] (I-ped2, II-ped 1) and [4] (III- Detection of Unusual Crowd Activity Outdoor, IV Detection of Unusual Crowd Activity-Indoor) for (a) normal activity scenes, and (b) abnormal activity scenes.

We assume that the values of ξ greater than or equal to 0.5 belong to the normal activity and is labelled as 1, and that less than 0.5 belong to the abnormal class and is labelled as 0. The same is the rule applied in the actual class labels that are used for the training. The use of the activation function in this manner helps in finding the dependencies which gives near optimal solutions on the class boundaries and provide a better response in the real-life applications when the evolved mathematical expression from GP training could be used as a generalized classifier.

The final fitness value is then computed by the summation of the XNOR (Exclusive NOR) operation between the actual class binary labels and the generated class binary labels. It is to be noted that the XNOR operation will generate a value of 0 for the false positive and the false negative cases, and a value of 1 for the true positive and the true negative cases. Hence, the fitness value will be higher for the candidates performing well towards the accuracy of the classifier generation buy generation.

4. Experimental Results and Discussion

In this section, we first describe the experimental settings which is followed by the results achieved. For the implementation of the proposed technique, we used AMD Ryzen 7, 4700U series processor of 2.0 GHz with Radeon Graphics and a system RAM of 8 GB. For the implementation, we used MATLAB programming environment with GBLAB [42] toolbox. In the training and the testing phases, we have used UCSD [33] and University of Minnesota [34] anomaly detection datasets. Some sample images from these datasets are shown in figure 3. The reason for choosing these datasets is that they have been extensively used by the research community and are considered some of the best available datasets for anomaly detection.

		Actual Classes	
		Normal Activity	Abnormal Activity
Predicted Classes	Normal Activity	TN	FN
	Abnormal Activity	FP	TP

Fig. 4. Confusion Matrix

For each of the candidate genome, a fitness score is calculated. We use the confusion matrix (figure 4), as is normally used in classification tasks, to detect the true positive rate (TPR) and the false positive rate (FPR).

Figure 5 shows the performance comparison of the proposed technique with some recent state of the art anomaly detection techniques using the UCSD dataset using the Area under the Curve (AUC) metric. The AUC refers to the area under the receiver operating characteristic (ROC) curve with the maximum possible value of 100%. It can be seen from the figure that the proposed technique outperforms most of the previous techniques. It is very near to the results achieved by Ye et al. [39] and Usman [31]. For the rest, it beats with a larger margin.

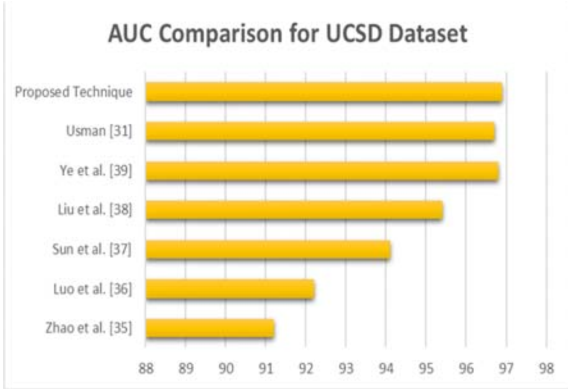

Fig. 5. Performance comparison of different techniques on UCSD dataset [33]

Figure 6 shows the performance comparison of the proposed technique with some recent state of the art anomaly detection techniques using the University of Minnesota dataset in terms of the ROC curve. Again, it can be seen that the proposed techniques perform better than most of the techniques including the social force model proposed by Mehran et al. [41] and spatio-temporal motion pattern model by Kratz and Nishino [28]. The performance is very near to the technique proposed by Usman [31] due to the fact that both techniques are similar in nature and are based on GP. The proposed technique performs a little better due to the utilization of and activation function in the training and the testing, and incorporating a new feature pertaining to motion-awareness. This helps the GP simulation to exploit the other hidden dependencies pertinent to the problem which were otherwise not explored.

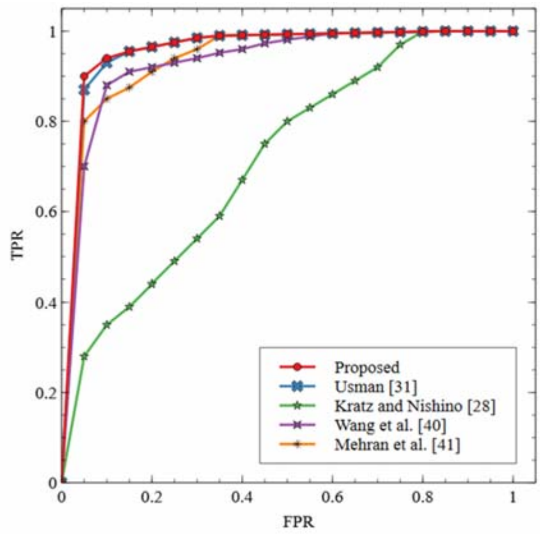

Fig. 6. Performance comparison of the proposed technique on University of Minnesota dataset in terms of ROC curve

Table 1: Performance comparison in Terms of Processing TIme with different techniques on University of minnesota dataset [34]

Method	Processing time
Proposed Method	0.16 s/frame
Usman [31]	0.15 s/frame
Nayan et al. [14]	0.18 s/frame
Wang et al. [40]	0.258 s/frame
Social Force Model [41]	> 1 s/frame

Table 1 presents the comparison in terms of the processing time. It is to be noted that the training time of the proposed technique is much higher due to the extensive GP based simulation and the evolution of the best mathematical expression. Once the best expression is evolved, it is generalized in nature and acts as a classifier which can then be directly used for real world applications. From the table, it can be seen that the proposed technique is a little slower compared to [31] due to the addition of the activation function step and an extra feature. We have also experimented with some of the activation functions during our simulation. The results of these functions are shown in table 2. We found that the *Sigmoid* function provided the best results. Nevertheless, the results obtained using the *step function* and the *Tanh* functions were also better than without using them.

Table 2: Performance comparison using different activations

Activation	Avg. AUC
Step Function	96.82%
Sigmoid	96.93%
Tanh	96.79%

Conclusion

In this paper, we introduce an automatic crowd video sequence anomaly detection system. For crowd motion, the proposed methodology uses a gradient-based approach with activation function and a motion aware feature. GP-based training simulation is used to evolve the best classifier in a stepwise enhancement process. Once evolved, the best mathematical expression is a classifier that performs better in exploiting the hidden dependencies in the decision space and is generic in nature. Experimental results validate the usefulness of the proposed technique and its superiority over the existing state of the art methods in terms of classification accuracy.

Acknowledgments

The authors highly acknowledge the research support provided by the Deanship of Scientific Research (DSR) at Saudi Electronic University, Riyadh, Saudi Arabia, through the approved project titled “Abnormal Crowd Behavior Detection for Time Varying Moving Sequences” with the research number 7661-CAI-2019-3-1-A.

References

[1] B. R. Kiran, D. M. Thomas, and R. Parakkal, “An overview of deep learning based methods for unsupervised and semi-supervised anomaly detection in videos," J. Imag., vol. 4, no. 2, p. 36, 2018.

[2] C. Dhiman and D. K. Vishwakarma, “A review of state-of-the-art techniques for abnormal human activity recognition,” Eng. Appl. Artif. Intell., vol. 77, pp. 21_45, Jan. 2018.

[3] A. B. Mabrouk and E. Zagrouba, “Abnormal behavior recognition for intelligent video surveillance systems: A review," Expert Syst. Appl., vol. 91, pp. 480_491, Jan. 2018.

[4] T. Wang, ``Abnormal event detection based on analysis of movement information of video sequence," Optik, vol. 152, pp. 50_60, Jan. 2018.

[5] W. G. Aguilar et al., “Statistical abnormal crowd behavior detection and simulation for real-time applications,” In International Conference on Intelligent Robotics and Applications, Springer, Cham., vol. 10463 LNAI, pp. 671–682, 2017, doi: 10.1007/978-3-319-65292-4_58.

[6] X. Zhang, Q. Zhang, S. Hu, C. Guo, and H. Yu, “Energy Level-Based Abnormal Crowd Behavior Detection,” Sensors, vol. 18, no. 2, Feb. 2018.

[7] M. Ravanbakhsh, M. Nabi, H. Mousavi, E. Sangineto, and N. Sebe ,“Plug-and-play CNN for crowd motion analysis: An application in abnormal event detection,” Proc. - 2018 IEEE Winter Conf. Appl. Comput. Vision, WACV 2018, vol. 2018-January, pp. 1689–1698, 2018, doi: 10.1109/WACV.2018.00188.

[8] L. Lazaridis, A. Dimou, and P. Daras, “Abnormal Behavior Detection in Crowded Scenes Using Density Heatmaps and Optical Flow,” 2018 26th Eur. Signal Process. Conf., pp. 2060–2064, 2018.

[9] S. Chaudhary, M. Aamir, and C. Bhatnagar, “Multiple Anomalous Activity Detection in Videos,” Procedia Comput. Sci., vol. 125, pp. 336–345, 2018, doi: 10.1016/j.procs.2017.12.045.

[10] W. Sultani, “Real-world Anomaly Detection in Surveillance Videos,” pp. 6479–6488.

[11] H. Wei, Y. Xiao, R. Li, and X. Liu, “Crowd abnormal detection using two-stream Fully Convolutional Neural Networks norm single frame norm optical flow norm,” 2018, doi:10.1109/ICMTMA. 2018.00087.

[12] M. Sabokrou, M. Fayyaz, M. Fathy, Z. Moayed, and R. Klette, “Deep- anomaly : Fully convolutional neural network for fast anomaly detection in crowded scenes Mohammad Sabokrou,” Comput. Vis. Image Underst., vol. 172, no. February, pp. 88–97, 2018, doi: 10.1016/j.cviu.2018.02.006.

[13] X. Zhang, Q. Zhang, S. Hu, C. Guo, and H. Yu, “Energy Level-Based Abnormal Crowd Behavior Detection,” Sensors, vol. 18, no. 2, Feb. 2018.

[14] N. Nayan, S. S. Sahu, and S. Kumar, “Detecting anomalous crowd behavior using correlation analysis of optical flow,” Signal, Image Video Process., vol. 13, no. 6, pp. 1233–1241, 2019, doi: 10.1007/s11760-019-01474-9.

[15] A. A. Afiq et al., “A review on classifying abnormal behavior in crowd scene,” J. Vis. Commun. Image Represent., vol. 58, pp. 285–303, 2019, doi: 10.1016/j.jvcir.2018.11.035.

[16] R. T. Ionescu, F. S. Khan, M. I. Georgescu, and L. Shao, “Object-centric auto-encoders and dummy anomalies for abnormal event detection in video,” arXiv, pp. 7842–7851, 2018.

[17] J. Wang and L. Xia, “Abnormal behavior detection in videos using deep learning,” Cluster Comput., vol. 22, no. s4, pp. 9229–9239, 2019, doi: 10.1007/s10586-018-2114-2.

[18] E. Duman and O. A. Erdem, “Anomaly Detection in Videos Using Optical Flow and Convolutional Autoencoder,” IEEE Access, vol. 7, pp. 183914–183923, 2019, doi: 10.1109/ACCESS.2019.2960654.

[19] Y. Liu, K. Hao, X. Tang, and T. Wang, “Abnormal Crowd Behavior Detection Based on Predictive Neural Network,” Proc. 2019 IEEE Int. Conf. Artif. Intell. Comput. Appl. ICAICA 2019, pp. 221–225, 2019, doi: 10.1109/ICAICA.2019.8873488.

[20] K. Pawar and V. Attar, “Deep learning approaches for video-based anomalous activity detection,” World Wide Web, vol. 22, no. 2, pp. 571–601, 2019, doi: 10.1007/s11280-018-0582-1.

[21] C. Direkoglu, “Abnormal Crowd Behavior Detection Using Motion Information Images and Convolutional Neural Networks,” IEEE Access, vol. 8, pp. 80408–80416, 2020, doi: 10.1109/ACCESS.2020.2990355.

[22] K. Singh, S. Rajora, D. K. Vishwakarma, G. Tripathi, S. Kumar, and G. S. Walia, “Crowd anomaly detection using Aggregation of Ensembles of fine-tuned ConvNets,” Neurocomputing, vol. 371, pp. 188–198, 2020, doi: 10.1016/j.neucom.2019.08.059.

[23] L. Xia and Z. Li, “A new method of abnormal behavior detection using LSTM network with temporal attention mechanism,” J. Supercomput., no. 0123456789, 2020, doi: 10.1007/s11227-020-03391-y.

[24] Z. Fan, J. Yin, Y. Song, and Z. Liu, “Real-time and accurate abnormal behavior detection in videos,” Mach. Vis. Appl., vol. 31, no. 7–8, pp. 1–13, 2020, doi: 10.1007/s00138-020-01111-3.

[25] R. Nawaratne, D. Alahakoon, D. De Silva, and X. Yu, “Spatiotemporal anomaly detection using deep learning for real-time video surveillance,” IEEE Trans. Ind. Informatics, vol. 16, no. 1, pp. 393–402, 2020, doi: 10.1109/TII.2019.2938527.

[26] G. Vallathan, A. John, C. Thirumalai, S. K. Mohan, G. Srivastava, and J. C. W. Lin, “Suspicious activity detection using deep learning in secure assisted living IoT environments,” J. Supercomput., no. 0123456789, 2020, doi: 10.1007/s11227-020-03387-8.

[27] B. Ramachandra and M. J. Jones, “Street Scene: A new dataset and evaluation protocol for video anomaly detection,” arXiv, pp. 2569–2578, 2019.

[28] L. Kratz and K. Nishino, “Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models,” in 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 1446–1453.

[29] S.Kullback and R.A.Leibler, "On Information and sufficiency," Ann.Math.Stat.,vol.22,no.1,pp.79-86, Mar. 1951.

[30] T. A. Myrvoll and F. K. Soong, “On divergence based clustering of normal distributions and its application to HMM adaptation,” EUROSPEECH 2003 - 8th Eur. Conf. Speech Commun. Technol., pp. 1517–1520, 2003.

[31] I. Usman, “Anomalous Crowd Behavior Detection in Time Varying Motion Sequences,” in 2019 4th World Conference on Complex Systems (WCCS), Apr. 2019, pp. 1–5, doi: 10.1109/ICoCS.2019.8930795.

[32] Y. Zhu, “Motion-Aware Feature for Improved Video Anomaly Detection,” p. 12.

[33] “UCSD Anomaly Detection Dataset.” http://www.svcl.ucsd.edu/projects/anomaly/dataset.html (accessed Apr. 16, 2021).

[34] “Monitoring Human Activity - Home,” Apr. 17, 2018. http://mha.cs.umn.edu/ (accessed Apr. 17, 2018).

[35] Y. Zhao, B. Deng, C. Shen, Y. Liu, H. Lu, and X.-S. Hua, “Spatio-Temporal AutoEncoder for Video Anomaly Detection,” in Proceedings of the 25th ACM international conference on Multimedia, New York, NY, USA, Oct. 2017, pp. 1933–1941, doi: 10.1145/3123266.3123451.

[36] W. Luo, W. Liu, and S. Gao, “A Revisit of Sparse Coding Based Anomaly Detection in Stacked RNN Framework,” in 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Oct. 2017, pp. 341–349, doi: 10.1109/ICCV.2017.45.

[37] Q. Sun, H. Liu, and T. Harada, “Online growing neural gas for anomaly detection in changing surveillance scenes,” Pattern Recognit., vol. 64, pp. 187–201, Apr. 2017, doi: 10.1016/j.patcog.2016.09.016.

[38] W. Liu, W. Luo, D. Lian, and S. Gao, “Future Frame Prediction for Anomaly Detection -- A New Baseline,” ArXiv171209867 Cs, Mar. 2018, Accessed: Apr. 19, 2021. [Online]. Available: http://arxiv.org/abs/1712.09867.

[39] M. Ye, X. Peng, W. Gan, W. Wu, and Y. Qiao, “AnoPCN: Video Anomaly Detection via Deep Predictive Coding Network,” in Proceedings of the 27th ACM International Conference on Multimedia, New York, NY, USA, Oct. 2019, pp. 1805–1813, doi: 10.1145/3343031.3350899.

[40] B. Wang, M. Ye, X. Li, F. Zhao, and J. Ding, "Abnormal crowd behavior detection using high-frequency and spatio-temporal features," Mach. Vis. Appl., vol. 23, no. 3, pp. 501–511, May 2012, doi: 10.1007/s00138-011-0341-0.

[41] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in 2009 IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2009, pp. 935–942, doi: 10.1109/CVPR.2009.5206641.

[42] "GPLAB - A Genetic Programming Toolbox for MATLAB." [Online]. Available: http://gplab.sourceforge.net/. [Accessed: 01-Apr-2019].

Imran Usman received his BE degree in Software Engineering from Foundation University, Pakistan in 2003 and MS Computer System Engineering from Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan in 2006. He joined Pakistan Institute of Engineering and Applied Sciences as a research scholar and received his PhD degree in 2010. From 2009 to 2010 he served at Iqra University Islamabad, Pakistan as an Assistant Professor in Department of Computing and Technology. From 2010 to 2012 he served as Assistant Professor and Senior In-charge Graduate Program in the Department of Electrical Engineering at COMSATS Institute of Information Technology Islamabad, Pakistan. He is presently serving as Associate Professor in College of Computing and Informatics, Saudi Electronic University, Kingdom of Saudi Arabia. His present research interests include machine learning, digital image processing, evolutionary computation and digital watermarking. Dr. Usman has a number of research papers to his credit and has supervised many BS, MS and PhD students. He is also a Senior Member of IEEE.

Abdulaziz A. Albesher Dr. Abdulaziz Al-Besher is the Dean of Computing and Informatics College at Saudi Electronic University. He is an Associate Professor of Information Technology, who holds a MSc in Information Systems Management and a PhD in Electronic and transformational government. He served in various IT positions in multiple institutions and currently he is a consultant in a number of public and private organizations. He has published various research papers in reputed journals and served as a reviewer in many international journals. He has also delivered special talks in several conferences and chaired various technical sessions. He has several achievement awards on Artificial Intelligence, Data Science and Leadership from different prestigious institutes. Dr. Abdulaziz's research focuses on innovation diffusion and adoption of technology. In addition, he is profoundly engrossed in the area of increasing productivity, sustainability and performance using several contemporary technologies, including Blockchain (in Finance & Healthcare), Internet of Things (IoT), Cyberdefense (using Artificial Intelligence, Machine Learning and Deep Learning).