

ABNORMAL EVENT DETECTION IN PEDESTRIAN PATHWAY USING GARCH MODEL AND MLP CLASSIFIER

Dr. Manjula Pattnaik,
Department of Accounting,
College of Business Administration,
Princess Nourahbint Abdulrahman University,
Riyadh, Kingdom of Saudi Arabia.
mpattnaik@pnu.edu.sa

Abstract: In computer vision, one of the complex research areas is video surveillance. It is very important to monitor the abnormal events in public places. Due to the technical advances, the usage of cameras is increased for surveillance purpose. As human operators are employed for the observation, their visual attention is reduced after long periods. Hence, an automated Abnormal Event Detection (AED) technique is designed in this study. It uses Generalized Autoregressive Conditional Heteroscedasticity (GARCH) which is a statistical model to model the events occurs in the pedestrian pathway. Before modeling, a series of preprocessing steps are employed to detect the moving objects. Multilayer Perceptron (MLP) is used to classify the parameters of GARCH model as normal event or abnormal event. Results show that the events are modeled by GARCH in an efficient manner which provides promising results for AED.

Keywords: Abnormal event detection, anomaly detection, GARCH modeling, multilayer perceptron, neural network.

I. INTRODUCTION

Human activity recognition based video anomaly classification and detection is discussed in [1]. At first, the input video is pre-processed by using kalman filtering. The moving objects are tracked by background subtraction. The actions of the human are also tracked. The classification of actions is made by Support Vector Machine (SVM) classifier. Anomaly detection in video surveillance is described in [2] using variable window size based on the size of the object. The object size and anomalous based features are extracted. The background subtraction, outlier detection and event modeling is used for the video anomaly detection.

The localization and detection of anomaly in surveillance videos is presented in [3]. The block based foreground segmentation is employed to analyze the moving objects in the video. The histogram of block-motion trajectories and clusters are extracted. The normal and abnormal events are detected by partition trajectory histogram. Buffer and anomaly detection based video storage algorithm is described in [4]. The buffer mechanism algorithm is used to store the information's in the video. It can analyze the anomalies in the stored video content only.

Trajectory simplification based traffic video anomaly detection is described in [5]. Initially, preprocessing is made by frame subtraction method. The anomalies are detected by the pseudo code framework. Then, the vehicles are tracked by the simple tracking algorithm. Multi feature extraction based framework for video anomaly detection is described in [6]. The input series are given to background model, foreground possibility, optical flow and structure of

interest response point. Then, the kinetic energy is applied. The location, velocity and appearance are used to identify the anomaly detection.

Automated video surveillance for crowd anomaly detection is described in [7]. The input video stream builds a video buffer. The spatio-temporal volume and texture, crowd region detection and wavelet transform is applied for decomposition. Finally, the video frames are compared and detected. Anomaly detection for traffic surveillance is described in [8]. At first, the frame acquisition is made from the input video. The background subtraction, morphological operations and object area are employed. The zones are detected by initializing the variables, main area and resetting the variables. The directions are detected and checked for the anomaly detection.

Sparse and low rank decompositions based video anomaly detection is described in [9]. Initially, the video frames are decomposed by principal component analysis and separated into two different sensors the low rank tensor and sparse tensor. The sparse tensor that consists of foreground pixels is used to detect the anomaly in the videos. Anomaly detection in surveillance videos is discussed in [10]. The spatial and temporal information's are given to the deep object detection and tracking. The information and appearance are given to the object detection network. The effective and scale invariant features like histogram variance of optical flow angle is used for video anomaly detection.

Overview of anomaly detection in video surveillance is discussed in [11]. The raw input data is given to event encoding for trajectory, motion labels, shape activity model and spatiotemporal model. The normal event and anomalous events are characterized by anomaly detector. Sub-trajectory based anomaly detection in video is described in [12]. Initially, the objects are detected in the input video. Then the detected objects are tracked in the object tracking method. The clustering based preprocessing method is used for anomaly detection.

An approach for AED by structural approach (SA) is discussed in [13]. The pedestrians in the frame are detected at first and then the individuals are described by structural context descriptor. The anomaly is detected by the variation in the structural context. The joint models of dynamics and appearance is discussed in [14] for AED. The expectation and maximization algorithm is used for learning the events.

In this study, GARCH model based AED system is presented using MLP classification. The rest of the paper is as follows: Section 2 gives the mathematical backgrounds of GARCH model and MLP classifier and how they are used for AED. Section 3 gives the results of AED system by GARCH and MLP. Section 4 concludes the AED system.

II. METHODS AND MATERIALS

The main objective of the AED system is the detection of abnormal event in the pedestrian pathway. Figure 1 shows the workflow of the AED system using GARCH and MLP classifier. It consists of three modules; preprocessing, GARCH modeling and MLP classification.

A. Preprocessing

In this phase, a series of operations are initially performed to detect the abnormal event effectively by the next phase. At first, the given video stream is converted into frames or images. Each frame is tested for moving objects based on the background subtraction. Let us consider, B be the background image which is

the only moving area taken from the training videos and F be the foreground image. The moving objects (M) in the frame F is defined by

$$M = \sum_{i=1}^M \sum_{j=1}^N F(i, j) - B(i, j) \quad (1)$$

where $M \times N$ is the size of the frame. The application of Eqn. (1) finds the moving objects in the frame. Then, simple thresholding is applied to M to get the binary image. It is defined by

$$B = \begin{cases} 0 & \text{if } M(i, j) = 0 \\ 1 & \text{else} \end{cases} \quad 1 \leq i \leq M, 1 \leq j \leq N \quad (2)$$

The connected component analysis is used to label the objects in the binary image. Based on the area occupied by the objects, small objects are removed.

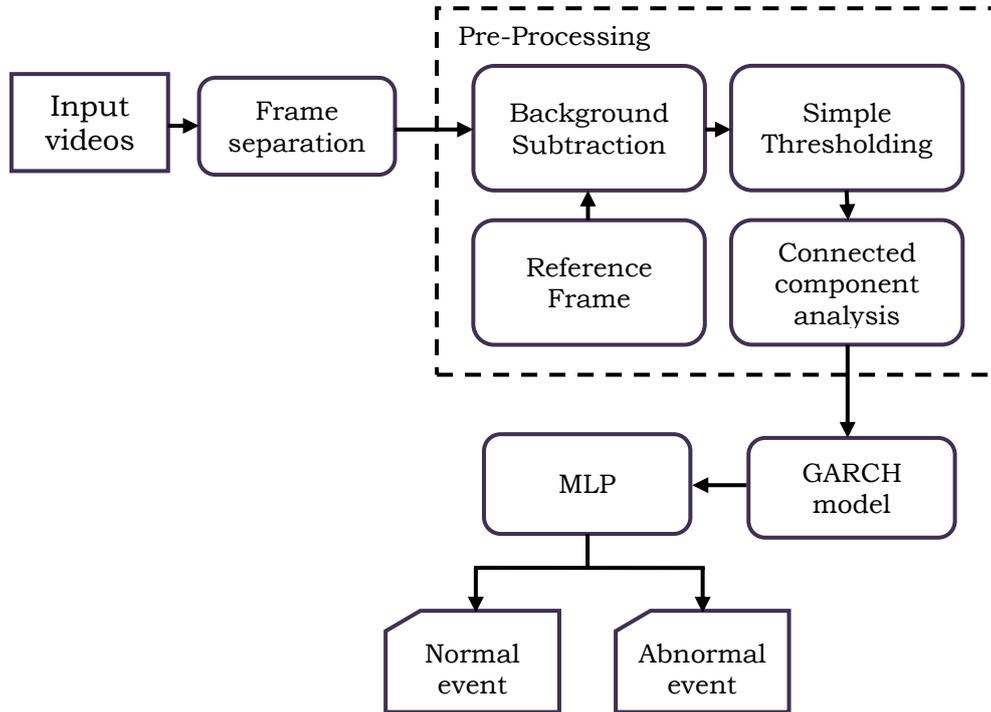


Fig. 1 Abnormal event detection in a pedestrian pathway

B. GARCH model

GARCH model overcomes the error variance of the autoregressive moving average model. Let us consider GARCH (s,t) model, where s and t are the order of δ^2 and τ^2 respectively. Then, the GARCH model is defined by [15],

$$x_k = y_k c + \tau_k \quad (1)$$

$$\delta_k^2 = \psi + \beta_1 \tau_{s-1}^2 + \dots + \beta_n \tau_{s-n}^2 + \alpha_t \tau_{s-t}^2 + \alpha_t \delta_{s-n}^2 = \Psi + \sum_{l=1}^n \beta_l \tau_{s-l}^2 + \sum_{l=1}^s \alpha_l \delta_{s-l}^2 \quad (2)$$

The heteroskedasticity test in economic models is the best white test in the time series to remove the GARCH errors. The exponential smoothing model is the alternative model for exponential weight moving average. It has more recent weight with the recent observations. The value of s is taken as 1. The length gap of the GARCH (s, t) model is established in three steps they are,

1. The best fitting is estimated AR(s) model

$$x_k = c_0 + c_1 x_{k-1} + \dots + c_q x_{k-q} + \delta_k = c_0 + \sum_{l=1}^s c_l x_{k-l} + \delta_k \quad (3)$$

2. Autocorrelations of τ^2 are computed and plotted.

$$\xi = \frac{\sum_{k=l+1}^P (\hat{\delta}_k^2 - \hat{\tau}_k^2)(\hat{\delta}_{k-1}^2 - \hat{\tau}_{k-1}^2)}{\sum_{k=1}^P (\hat{\delta}_k^2 - \hat{\tau}_k^2)^2} \quad (4)$$

3. The standard deviation of $\xi(l)$ is $1/\sqrt{P}$. The GARCH errors are less than the individual values. The number of lags is estimated by the Ljung-Box test. It follows the Z^2 distributions with the squared residuals of τ_k^2 are correlated. In this study, GARCH (1,1) model is used to extract the parameters from the events.

C. MLP Classifier

MLP is a class belongs to feed forward artificial neural network. It is referred as the neural network composed of multiple layers of perceptrons. It consist of three layers namely, input, hidden and output layers. Figure 2 shows the architecture of MLP.

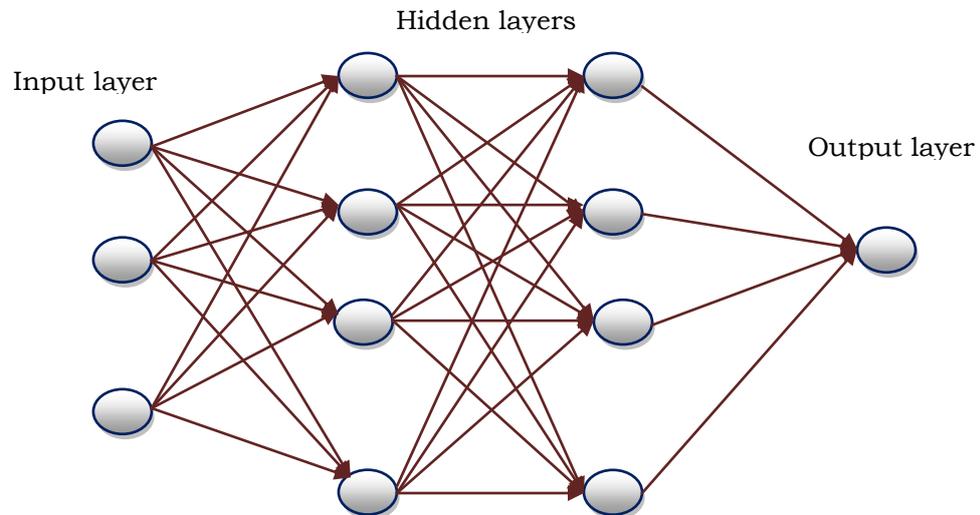


Fig. 2 MLP architecture

The non-linear activation function is used for the each node except input nodes. It uses the back propagation for the training. MLP is distinguished into linear perceptron using non-linear activation and multiple layers. The non-linearly separable data are differentiated by the MLP. All neurons have the linear

activation function which maps the input and output weights of each neuron. The nonlinear activation function has developed the model to the frequency of biological neurons, potential action and firing. The rectified linear unit is the activation function which is frequently used. The activation function of MLP is sigmoid. The sigmoid is defined by,

$$z(k_j) = \left(1 + e^{-k_j}\right)^{-1} \quad (5)$$

where Z_j is the output of the j^{th} node and k_j is the weight of the input of the connections. As the system classifies the given event as either normal or abnormal event, a pure linear function is used in the output layer.

III. RESULTS AND DISCUSSION

The performance of AED system in a pedestrian pathway is discussed in this section. It uses the video sequences obtained from University of California San Diego (UCSD) database. The crowd densities in pedestrian pathway vary from sparse to heavy with one or more abnormal events. Table 1 shows the description about UCSD database.

TABLE 1 UCSD database description

Database	#training videos	#testing videos	#frames per video	#normal frames	#abnormal frames	Frame size
PED1	34	36	200	5000	3400	158 x 238 pixels
PED2	16	12	200	2566	2384	320 x 240 pixels

The difference between the videos in PED1 and PED2 is the position of the camera that captures the movement of pedestrian. The pedestrian movement in the later one is parallel to the camera plane and the former one is towards and away from the camera. A binary flag in the ground truth indicates whether a frame has abnormal event such as bikers, skaters and small cars or not. Figure 3 shows some abnormal events in a pedestrian pathway.

In this study, abnormal event in a video clip is detected in a frame level manner. In the classification or abnormal event detection phase, if there is any abnormality found then the frame is classified as abnormal immediately and the process skip to next frame for further detection. To analyze the performance of AED system, at first False Positive Rate (FPR) and True Positive Rate (TPR) are calculated. They are defined by

$$TPR = \frac{\text{Number of abnormal frames correctly classified as abnormal}}{\text{total number of frames tested}} \times 100$$

$$FPR = \frac{\text{Number of abnormal frames correctly classified as normal}}{\text{total number of frames tested}} \times 100$$

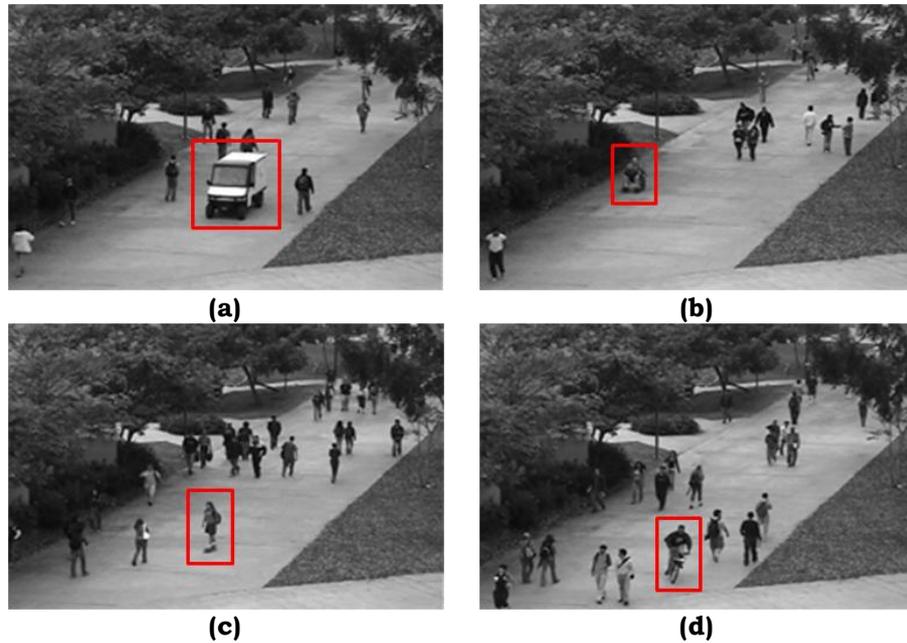


Fig. 3 Abnormal events in UCSD- PED1 database (a) cart (b) wheel chair (c) skater (d) biker

From these two performance metrics, Receiver Operating Characteristics (ROC) curve is drawn for further analysis. The performance of AED system by GARCH and MLP is compared with MDT [13] and SA [14]. Figure 4 and 5 show the obtained ROC curve for UCSD-PED1 and UCSD-PED2.

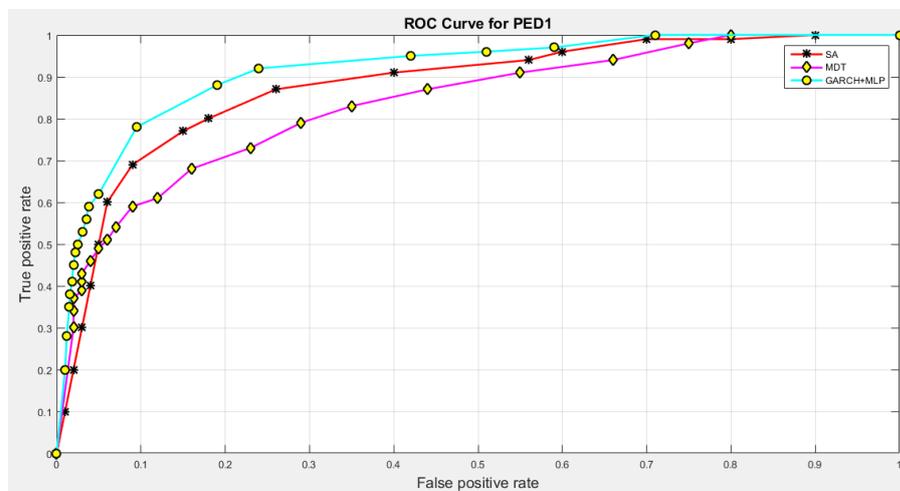


Fig. 4 ROC curve using UCSD- PED1 database

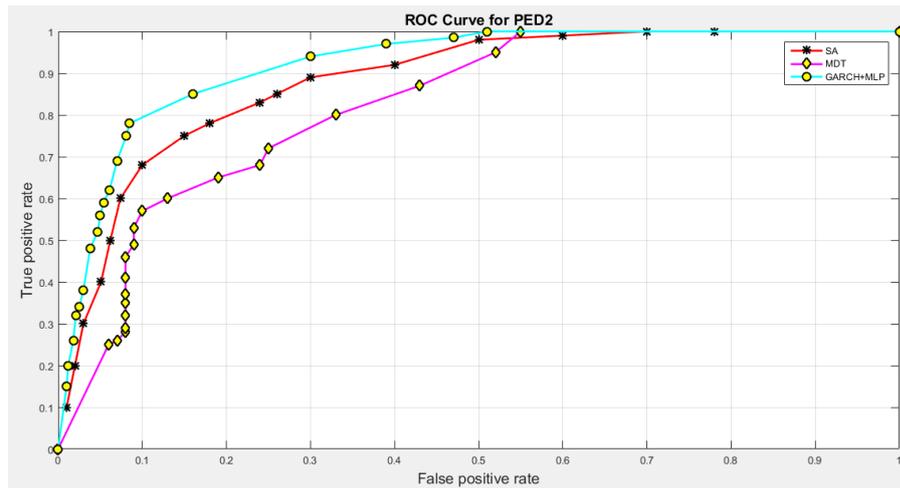


Fig. 5 ROC curve using UCSD- PED2 database

It is observed from the Fig. 4 and 5 that the AED system using GARCH and MLP gives better performance as the ROC curve for UCSD-PED1 and UCSD-PED2 are above SA and MDT techniques. At 0.1 FPR, the TPR of GARCH+MLP approach for UCSD-PED1 and UCSD-PED2 are 0.80 which outperforms all approaches.

IV. CONCLUSION

In this study, an efficient AED system is designed for video surveillance. It uses GARCH statistical model to model the events occur in the pedestrian pathway. All the events such as person, cart, wheel chair, skater and biker are detected by a series of preprocessing steps such as frame separation, motion detection and connected component analysis. The obtained GARCH model parameters are fed to the MLP classifier which acts as a binary classifier to classify the events as normal or abnormal. The system uses frame level detection which means that if a abnormal event is detected in the frame, then the frame is classified as abnormal. Results shows that the AED system provides better result when model the events using GARCH and MLP classifier.

REFERENCES

- [1]. S. Zaidi, B. Jagadeesh, K.V. Sudheesh, and A.A. Audre, "Video Anomaly Detection and Classification for Human Activity Recognition", International Conference on Current Trends in Computer, Electrical, Electronics and Communication, 2017, pp. 544-548.
- [2]. J.C. SanMiguel, J.M. Martinez, and L. Caro-Campos, "Object-size invariant anomaly detection in video-surveillance", International Carnahan Conference on Security Technology, 2017, pp. 1-6.
- [3]. X. Zang, G. Li, Z. Li, N. Li, and W. Wang, "An object-aware anomaly detection and localization in surveillance videos", International Conference on Multimedia Big Data, 2016, pp. 113-116.

- [4]. J.N. Lv, X.Z. Li, Z.J. Zhang, G.H. Gao, and S.P. Wang, "A Video Storage Algorithm Based on Buffer and Anomaly Detection", International Conference of Information Science and Management Engineering, 2010, pp. 442-445.
- [5]. M. Isaloo, and Z. Azimifar, "Anomaly detection on traffic videos based on trajectory simplification", Iranian Conference on Machine Vision and Image Processing, 2013, pp. 200-203.
- [6]. Q. Li, and W. Li, "A novel framework for anomaly detection in video surveillance using multi-feature extraction", international symposium on computational intelligence and design, 2016, pp. 455-459.
- [7]. J. Wang, and Z. Xu, "Crowd anomaly detection for automated video surveillance", International Conference on Imaging for Crime Prevention and Detection, 2015, pp. 1-6.
- [8]. S. Shashikar, and V. Upadhyaya, "Traffic surveillance and anomaly detection using image processing", International Conference on Image Information Processing, 2017, pp. 1-6.
- [9]. L. Tran, C. Navasca, and J. Luo, "Video detection anomaly via low-rank and sparse decompositions", Western New York Image Processing Workshop, 2012, pp. 17-20.
- [10]. X. Li, W. Li, B. Liu, Q. Liu, and N. Yu, "Object-Oriented Anomaly Detection in Surveillance Videos", International Conference on Acoustics, Speech and Signal Processing, 2018, pp. 1907-1911.
- [11]. P. Kaur, M. Gangadharappa, and S. Gautam, "An Overview of Anomaly Detection in Video Surveillance", International Conference on Advances in Computing, Communication Control and Networking, 2018, pp. 607-614.
- [12]. D.V. Ngo, N.T. Do, and L.A.T. Nguyen, "Anomaly detection in video surveillance: A novel approach based on sub-trajectory", International Conference on Electronics, Information, and Communications, 2016, pp. 1-4.
- [13]. V. Mahadevan, W. Li, V. Bhalodia and N. Vasconcelos, "Anomaly detection in crowded scenes", IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 1975-1981.
- [14]. Y. Yuan, J. Fang and Q. Wang, "Online anomaly detection in crowd scenes via structure analysis", IEEE Transactions on Cybernetics, Vol. 45, No. 3, 2005, pp. 548-561.
- [15]. T. Bollerslev, "Generalized Autoregressive Conditional Heteroskedasticity", Journal of Econometrics, Vol. 31, No. 3, 1986, pp. 307-327.