Proceedings of the Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021).
IEEE Xplore Part Number: CFP21ONG-ART; 978-0-7381-1183-4

2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) | 978-1-6654-1960-4/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICICV50876.2021.9388587

UNUSUAL ACTIVITY AND ANOMALY DETECTION IN SURVEILLANCE USING GMM-KNN MODEL

G. Shanmuga Priya[1], M.Latha[2] , K.Manoj[3], Siva Prakash[4]
[1,3,4] Student, [2] Assistant Professor
Computer Science and Engineering
Vels Institute of Science, Technology & Advanced Studies (VISTAS)
Chennai, Tamil Nadu, India

Abstract- The major purpose of surveillance is to keep track of a region for security enhancement. The traditional system is being accustomed so far, where it uses a lot of manpower and energy to manually monitor the anomalous events. It therefore results in a waste of time and resources. This paper is a part of intelligent security systems to detect any unwanted happenings or anomalies in a locale using surveillance. An approach to detect unusual activity in surveillance using GMM-KNN model has been proposed. The dataset is acquired from the industrial premises of NLC India Limited. The project proposal has the following stages. The elicitation of frames is performed from the obtained dataset. The initial foreground detection is accomplished by using the GMM. The post processing step to neglect unwanted noises and objects is done using filters and morphological operations. Statistical features are extracted for each object to detect anomaly. Feature based Detected foreground objects are classified into anomaly and background based on KNN classifier.

Keywords - Anomalies, Background Subtraction, Foreground detection, Gaussian mixture model (GMM) and K nearest neighbor (KNN), Surveillance.

I. INTRODUCTION

GMM-KNN based anomaly detection system can be used in various fields of industrial premises, traffic surveillance, malls and any crowded areas to detect obstacles and anomalies. By subtracting the pure background and the current image using background subtraction the moving foreground is captured.[1] The previously used conventional method failed to detect an object in a dynamically changing environment. It couldn't accurately describe the dynamic objects such as shaking leaves. Image processing, pattern recognition and computer vision are broad fields of research concerning the analysis of image processing and video compression [2][24]. Profiting from the development in those field, automated video surveillance has developed as a research topic that has gained a lot of popularity. Due to the insecure environment, surveillance cameras have been equipped in various places to monitor the day to day activities. With the videos being captured round and round by the cameras and due to the enormous amount of data being generated every now and then, it becomes tedious to monitor them manually. By doing so there is also a wastage of manpower and energy.

In the last few decades, video surveillance systems has experienced a tremendous development , most denotable after the major attacks such as 9/11 attacks on twin towers in new York , Terrorist bomb explosions in spain on march 11,2004 and the Pulwama Kashmir attack in June 2019.Not only this but several major attacks and disastrous events insisted on the development of the intelligent surveillance system. The rapid growth of video surveillance in turned produced an enormous amount of data which resulted in the workload for CCTV operators to detect anomalies. To alleviate this problem automatic video analysis procedures came into existence. A recent statement by USA is that it has announced a $2.9 million package to help India as a part of US global assistance initiative to help with the case management and surveillance activities. The government of all major developing countries are investing a huge amount in surveillance system to ensure a stronger security. With the diversified real-world anomalies, which are complicated, it becomes difficult to list every possible event [3].

Machin learning algorithms can be developed with the training set along with its attributes which provides information about the anomaly types [4]. An algorithm for 2D median filtering based on storing and updating the gray level histogram has been implemented which showed faster execution than traditional methods [5]. Sparse coding-based tactics is considered as a characteristic method to achieve the desired results. In this approach the dataset usually contains only a small portion of normal actions, the anomalies cannot be reproduced from the normal actions. And in the drastically changing diverse environments in surveillance cameras, these methods produce unwanted noises, poor accuracy and high untrue alarm frequency [6].

II. RELATED WORKS

In the field of motion detection, Background Subtraction (BS) has been an effectual way of separating the moving objects in a static recorded video sequence. BS is mainly based on two perspectives. The False Negatives (FN) and False Positives (FP). Toyama et al. [13] and Panahi et al. [14] executed various video trails for pixel-based BS algorithms. The FN and FP are measures dependent to each other, where FN increases FP decreases and vice versa. This paper suffers with a limitation of a single threshold, where a single {FN,FP} may not be sufficient. Further Chalidabhongse et al. [15] proposed another method for BS based on the analysis for disturbances. In the initial stage, the limitation in previous works were fixed by adjusting ad-hoc threshold. The vector disturbances in all the directions of RGB space corrupts the background of each video trails. Since the pixel distribution is usually a unimodal, this proposal fails with a multibackground environment that involves stimulated foreground objects that are moving. It does not involve any postprocessing steps or region-based classification.

In the case of target detection by P. Ribeiro et al. [16] which involves car or any object in a localization, the motion detection method focuses on objects and club together the foreground pixels together as globules. The position of these blobs is used for the analysis[26]. In the case where a moving object obstruct another object, it is connected as a huge single moving globule. In our proposal this limitation is overcome by a high-end postprocessing step. The Gaussian Mixture Model is used in our proposal involves a multi dynamic environment that has vibrant textures such as shaking leaves, trees etc.,) Stauffer and Grimson [17] proposed the use of multimodal ML models. The probability of appearances of a colour at a pixel can be given by GMM model. Once every Gaussian is updated, the weights are normalized, and k distribution is ordered, and the background is chosen with a threshold. And then the pixels based on the distances are labeled based on the deviations. The proposal was further motivated by the improved for real time tracking with the help of shadow detection by P. KaewTrakulPOng et. al,[18]. It worked with low quality videos with shadows or filming dark areas, unwanted distortion and noises occurs. The noises are processed using the morphological operation, mean and median filters. NN classifier uses the feature vector and TCNN, for introducing pooling layer. Regarding image classification and object detection deep learning has been implemented. But its impact is limited due to video data, whereas CNN uses frame level concepts [19][27].

For differentiating and identifying various patterns and unusual patterns, outliers were used. While using video surveillance variety of anomalies can be captured and detected via labelling, processing and detecting. The anomaly detected in this paper gave an accuracy of 98.5% [20]. Anomaly detection for vehicle and human behavior used NN and machine learning methods. This paper used R-CNN for identifying objects and its scene location as its first step and using the optical flow [21]. Regarding image classification and object detection deep learning has been implemented. But its impact is limited due to video data, whereas CNN uses frame level concepts [22][25]. Detection from video records usually involves more traffic and erroneous things due to datasets. YOLO (You Only Look Once) has been implemented for improving the speed and accuracy at 40 frames per second [23].

Dr. Dorothy Dening,1985 introduced a system to analyze the audits of government and track the events. This system was named as Intrusion Detection Expert System . It was the first step into IDS technology. Dr. Dorothy Dening,1986 The first intrusion detection model was formulated in this year, he developed the model with six main components Subject, object, anomaly record, anomaly rules, profiles and audit lists Dr. Dorothy Denning,1987 Behavioral analysis approach was presented in this paper, behavior analysis looks for a certain behavior and deviation [28].

Teng, Chen, And Lu, 1990 This paper induces time based inductive systems to capture a behavior pattern. The temporal sequence patterns were represented in the form of sequential patterns. It used a logical inference known as inductive generalization [29].

.

III. PROPOSED WORK

Although the previous methods to detect anomaly may be appealing, they are grounded on a static environment. In case of a dynamic environment where the object and the background environment keep dynamically changing, the real challenge of object detection happens [4][11]. Axelsson, S., 1999[30], Axelsson proposed the IDE in LAN networks. The connectivity, response compatibility, intrusion detection techniques and vulnerabilities were discussed. Lane and Brodley 1998[31] introduced an entity known as applied instance learning and it learns from the temporal data. Based on the relation to stored instances a new instance has been classified. Lee W. and Stolfo S. and Mok K, 1999[32] extracted the features using audit programs and the framework for Intrusion Detection was proposed. Outlier detection is a rapidly growing topic in the field of research and development. An unnoticed outlier is the major reason for most of the disasters. In the field of computer vision, the major component of outlier or anomaly detection is pattern recognition. Any event that deviates from the determined pattern is term to be an anomaly. There are several attempts to detect an anomaly in video surveillance [7].

In this paper, an anomaly detection method has been proposed in a constantly changing environment using the Gaussian Mixture Model and by labeling the K Nearest Neighbor. The video is trained in such a way that the normal and anomalous events are taken as positive and negative.

The Proposed method can be summarized on nine stages : (i) Acquisition of Video Dataset (ii) Elicitation of the video into frames (iii) Foreground detection initialization using the GMM (iv) Post processing of images (v) Final foreground Detection (vi) Statistical Feature Extraction and Labelling (viii) Classification using K Nearest Neighbour (ix) Anomaly Detection. If an anomaly is detected, an alarm is generated by the system as the last and final step

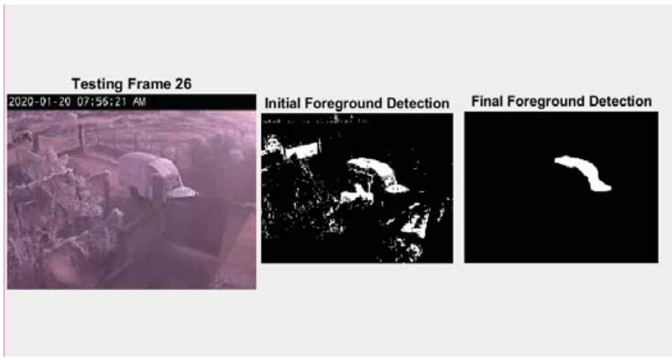
- Elicitation of frames from the acquired dataset. The initial foreground object Detection, where GMM is implemented.
- The post processing step to neglect the small objects unwanted noise, shaking tress, leaves etc., using median Filter, mean filter and morphological Operation.
- The blob analysis is used to generate the bounding box to each foreground object detected in the training frame. The statistical features are detected and labelled and are extracted to detect anomaly. The feature based Detected foreground objects are classified into anomaly and background based on KNN classifier
- Our proposed has a superior accuracy and performance as compared to the conventional anomaly detection approaches. The rate of classification accuracy and the computational time adds advantage to our proposal.

Dataset

In order to improve performance our BS proposal has been tested on various instance of the video frames. Our Dataset is a privately obtained dataset from the industrial premises of NLC India Limited. It is composed of video sequence of a restricted locale which has a dynamic environment with different textures. The real time video is divided into 10,000 frames containing the training, testing and validation data. They are given as input value for the GMM-KNN model. The precise ground truth value and the statistical values for each frame is being calculated and the blob analysis for enhanced detection. Further the proposal has been tested on various dataset downloaded from the online source for enhance anomaly detection purposes.

Background Subtraction

In image processing and Computer Vision Background subtraction otherwise known as Foreground Detection is mainly carried out, during classification of an image in a video. In our application environmental details is not needed; rather the information changes in the scene is required. This method is used for object localization where it separates the foreground from the background which is shown in fig 3.1.



3.1 Background image and Foreground Subtraction

In the surveillance system wherein the cameras are placed stationary . The quality of the video alters when the environment keeps changing drastically and it becomes difficult and challenging to analyze the video.

A perfect background subtraction model must be able to

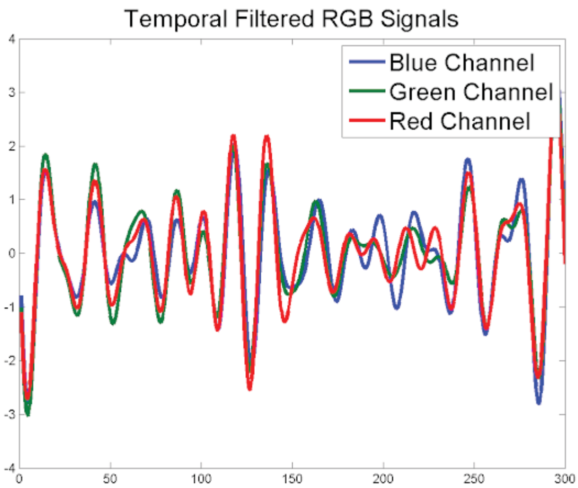
- Be strong to changes like waves, leaves, trees and other repetitive movements
- Have a high accuracy and develop an estimated model

- Produce less false alarm rate

The basic approach is detecting the moving object from the current frame. Some Conventional methods of Background Subtraction are:

Temporal Average Filter

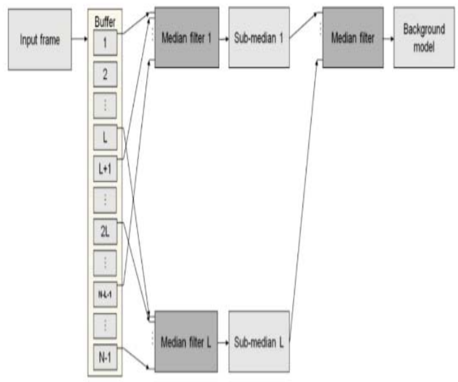
- In temporal Average Filter the pixel values are compared. In modeling the background, all the images are examined with a period called the training period.
- The median is calculated pixel by pixel. If the pixel matches the background, the value is included or else it is classified as a foreground and has been depicted in fig 3.2.
- It is not so efficient as it does not contain laborious statistical base and it requires a high computational speed [5]



3.2 Temporal Filter

Median Model

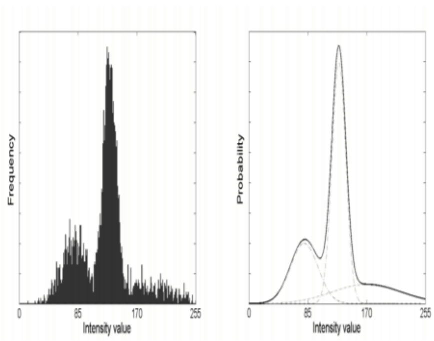
For modelling the background pixels in a video median value of N frames can be used without degrading foreground moving objects. It makes the background visible for frames higher than N/2. A buffer is used to store the last N frames and its flow diagram is shown in fig 3.3.



3.3 Hierarchical median approach

Gaussian Mixture Model

Most of the models use a single description for the background. Gaussian Mixture Models (GMMs) can be used for complex distributions [8]. This approach classifies each pixel using a model which consists of the combination of different classes. Figure 3.4 shows the empirical distribution of the intensity values.



3.4 Gaussian Mixture Model.

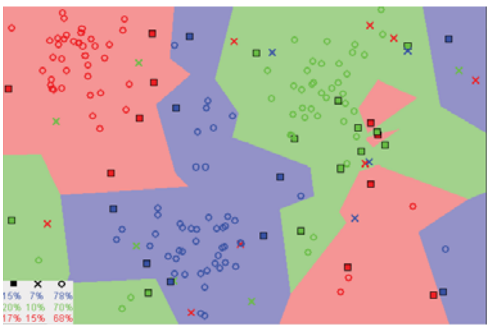
Gaussian Mixture Models: Algorithm

- First, compute the sample means of each class in the training data
- Use fitgmdist function for total training data with regularization parameter and a number of mixtures
- Number of mixture is always in multiples, the number of classes corresponding to data sets respectively
- Regularization parameter ensures estimated covariance matrices
- Then, find the smallest distance between each class sample mean and each mixture
- The class associated with minimum distance is assigned to each mixture

- The cluster function is used and helps to Count the incorrect classifications
- The probability of Error can be checked by dividing the number of incorrect class assignments to number of test vectors

K nearest neighbor

It is a non-parametric object classification method where the object is classified based on the k closest elements. It is easy to be implemented in supervised machine learning for classification problems.[7][9][11]. A supervised machine learning as the name implies it is under supervision and completely depends on labeled data.[10][12]. It generates an appropriate output when an unlabeled data is tested. The KNN algorithm identical things are labeled close to each other. Similar datapoints that exists close to each other. By calculating the distance between two points, the graph is plotted and is shown in fig 3.5.



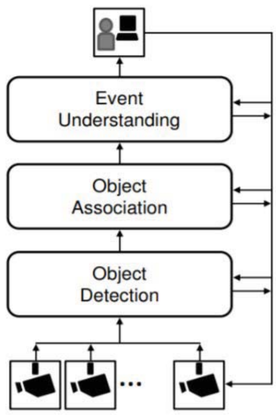
3.5 KNN Model

KNN Algorithm

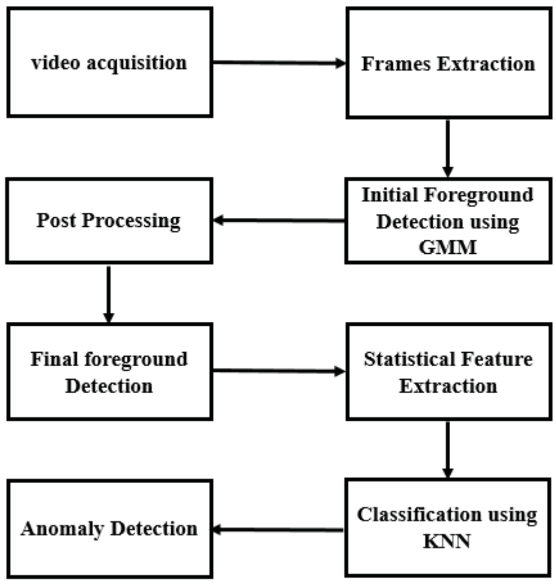
- Load the data
- K is initialized to the selected number of neighbours
- Calculate the distance between the queries
- Add the index and the distances to the collection in order
- ascend the collections of indices and distances in order [10]
- The first K entry is picked from the collection
- Obtain the labels of the chosen k entries
- If regression occurs, the mean of k label is returned
- If classification occurs, return the mode of k labels

Architecture and Data Flow Diagram

The general flow diagram of video surveillance starts with event happening, its object and detecting them. Its has been depicted in fig 3.6. The proposed architecture diagram starts with extracting the videos with frames, identifying its foreground and background object detection, feature extraction and classification using KNN. The proposed flow diagram has been shown in fig 3.7.



3.6 Video Surveillance General Flow



3.7 Proposed Architecture

IV RESULTS AND DISCUSSION

In the process of video analytics for detecting anomalies, the input images are trained using GMM KNN model to obtain the desired result. Noise during the reconstruction phase has been a great challenge for almost all researches. In this proposal the noise has been resolved using mean, median and morphological operations. The proposed method has shown a greater computational efficiency, the computational time is less, and the rate of classification accuracy is high. This experimental method has obtained an accuracy of 97 percent which is comparatively higher than the previous proposed methods. The accuracy comparison of various methods is represented below in table 4.1. Various screenshots of frame extraction and detection has been shown in fig 4.1,4.2,4.3,4.4.

Table 4.1 Accuracy of Existing and Proposed Methods

Methods	Accuracy rate
Deep Residual Networks Lucas A. Thomaz et.al,	**87.9 %**
K means Clustering Algorithm shivam et.al,	**90.5%**
Real time anomaly detection Abdullah Ghanj et.al	**95.4%**
GMM-KNN model (Proposed Method)	**97%**

Fig. 4.1 Normal frame

Fig. 4.2 Anomalous Frame

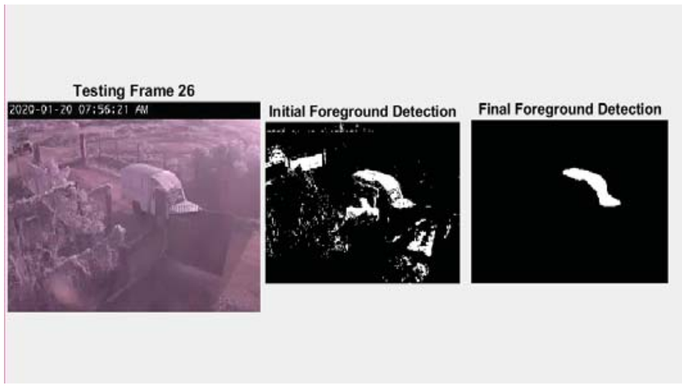


Fig. 4.3 Initial and foreground Detection

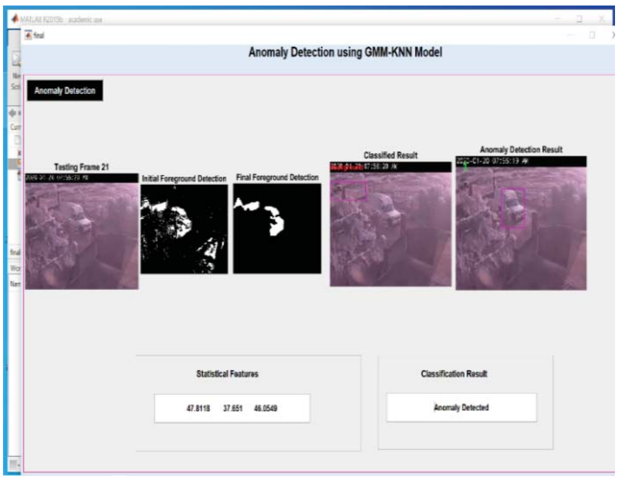


Fig. 4.4 Anomaly Detection

V CONCLUSION

In this paper, background subtraction model has been proposed to detect anomalous events in a locale. This method improvised the computational speed, accuracy and detects motion on all kinds of environment. GMM has been used for the background detection and K nearest neighbour for clustering. A major challenge of reconstruction noise has been cleared in this proposal and it further reduces the occurrence of false alarm rate. The proposed model requires very minimal adaption and can be easily implemented and an accuracy rate of 97% has been achieved. In a secured environment like NLC, this method can be used which helps to detect anomalies in an efficient way. A video surveillance camera can detect the unusual objects and give an alarm message which helps in preventing further mis happenings in various secured places. By using the background subtraction model the foreground object is detected. Noise during the reconstruction phase has been a great challenge for almost all researches. In this proposal the noise has been resolved using mean, median and morphological operations. The Experimental results has shown a greater computational efficiency, the computational time is less, and the rate of classification accuracy is high.

REFERENCES

[1] Piccardi, M. (2004). "Background subtraction techniques: A review". IEEE International Conference on Systems, Man and Cybernetics, pp.3099-3104.

[2] V. Saligrama Y. Benezeth, P. Jodoin and C. Rosenberger. 2009," Abnormal events detection based on spatio-temporal co-occurrences", in Proc. IEEE Conference Computer Vision Pattern Recognition, Workshops, Jun.2025, 2458-2465.

[3] Waqas Sultani, Chen Chen, and Mubarak Shah, "Real-world anomaly detection in surveillance videos", in CVPR, 2018.

[4] Anishchenko, Lesya. (2018), "Machine learning in video surveillance for fall detection" pp:99-102. 10.1109/USBEREIT.2018.8384560.

[5] Huang, Thomas S.; Yang, George J.; Tang, Gregory Y."A fast-two-dimensional median filtering algorithm" (PDF). IEEE Transactions on Acoustics, Speech, and Signal Processing. **27** (1): 13–18.

[6] Rohankar, Jayant (Nov 2013), "Survey on Various Noises and Techniques for Denoising the Color Image", International Journal of Application or Innovation in Engineering & Management(11), Retrieved 15 May 2015.

[7] Hawkins, Douglas M," Identification of outliers", Vol. 11. London: Chapman and Hall, 1980.

[8] Zivkovic Z," Improved Adaptive Gaussian Mixture Model for Background Subtraction", International Conference on Pattern Recognition, ICPR, vol. 2, pp. 28–31 (2004)Google Scholar

[9] T. Sandhan, A. Sethi, T. Srivastava and J. Y. Choi, "Unsupervised learning approach for abnormal event detection in surveillance video by revealing infrequent patterns," 2013, 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013), Wellington, 2013, pp. 494-499

[10] Laskov, Pavel & Düssel, Patrick & Schäfer, Christin & Rieck, Konrad. (2005),"Learning Intrusion Detection: Supervised or Unsupervised?".

[11] Coomans, Danny; Massart, Desire L. (1982). "Alternative K-Nearest Neighbour rules in supervised pattern recognition: Part 1. k-Nearest neighbour classification by using alternative voting rules", Analytica Chimica Acta. 136: 15–27.

[12] Srivastava, Nitish, Elman Mansimov, and Ruslan Salakhudinov," Unsupervised learning of video representations using lstms", International conference on machine learning. 2015.

[13] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers," Wallflower: principles and practice of background maintenance", International Conference on Computer Vision, 1:255–261, 1999.

[14] S. Panahi, S. Sheikhi, S. Hadadan, and N. Gheissari, "Evaluation of background subtraction methods. International conference on Digital Image Computing: Techniques and Applications", pages 357–364, 2008.

[15] T.H. Chalidabhongse, K. Kim, D. Harwood, and L. Davis, "A perturbation method for evaluating background subtraction algorithms", International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2003.

[16] D. Hall, J. Nascimento, P. Ribeiro, E. Andrade, P. Moreno, S. Pesnel, T. List, R. Emonet, R.B. Fisher, J.S. Victor, and J.L. Crowley, " Comparison of target detection algorithms using adaptive background models", International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, pages 113–120, 2005.

[17] C. Stauffer and W.E.L. Grimson, "Adaptive background mixture models for real-time tracking", International conference on Computer Vision and Pattern Recognition,1999.

[18] P. KaewTraKulPong and R. Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection", Workshop on Advanced Video-based Surveillance Systems conference, 2001.

[19] R. Hou, C. Chen, and M. Shah, "Tube convolutional neural network (t-cnn) for action detection in videos", in ICCV, 2017

[20] R. J. Franklin, Mohana and V. Dabbagol, "Anomaly Detection in Videos for Video Surveillance Applications using Neural Networks," 2020, Fourth International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 2020, pp. 632-637, doi: 10.1109/ICISC47916.2020.9171212

[21] Mohammad Farhadi et al., "AAD: Adaptive Anomaly Detection through traffic Surveillance videos", arXiv:1808.10044, 2018.

[22] M R Nehashree et al., "Simulation and Performance Analysis of Feature Extraction and Matching Algorithms for Image Processing Applications", International Conference on Intelligent Sustainable Systems (ICISS-2019).

[23] Arka Prava Jana et al., "YOLO based Detection and Classification of Objects in video records", International Conference On Recent Trends In Electronics Information Communication Technology(RTEICT), 2018.

[24] Yong Shean Chong and Yong Haur Tay. 2017. Abnormal Event Detection in Videos Using Spatiotemporal Autoencoder. Springer International Publishing, 189-196.

[25] Chalapathy, Raghavendra, and Sanjay Chawla. "Deep Learning for Anomaly Detection: A Survey." arXiv preprint arXiv:1901.03407 (2019)

[26] Landi, Federico, Cees GM Snoek, and Rita Cucchiara. "Anomaly Locality in Video Surveillance." arXiv preprint arXiv:1901.10364 (2019)

[27] Sun, Jiayu, Jie Shao, and Chengkun He. "Abnormal event detection for video surveillance using deep one-class learning." Multimedia Tools and Applications 78.3 (2019): 3633-3647

[28] Dr. Dorothy Dening,1985, "An Intrusion – Detection Model", IEEE transactions on software engineering, vol,SE-13, No:2, Feb 1987, 222-232.

[29] H. S. Teng, K. Chen and S. C. Lu, "Adaptive real-time anomaly detection using inductively generated sequential patterns," Proceedings. 1990 IEEE Computer Society Symposium on Research in Security and Privacy, Oakland, CA, USA, 1990, pp. 278-284, doi: 10.1109/RISP.1990.63857

[30] Stefan Axelsson,"The base-rate fallacy and its implications for the difficulty of intrusion detection ", Proceedings of the 6th ACM conference on Computer and communications security, November 1999, Pages 1–7, https://doi.org/10.1145/319709.319710

[31] Terran Lane and Carla E. Brodley. 1998,"Temporal sequence learning and data reduction for anomaly detection" in Proceedings of the 5th ACM conference on Computer and communications security (CCS '98). Association for Computing Machinery, New York, NY, USA, 150–158. DOI:https://doi.org/10.1145/288090.288122

[32] Lee, Wenke & Stolfo, Salvatore & Mok, Kui. (1999),"A data mining framework for building intrusion detection models", Security and Privacy, 1999. Proceedings of the 1999 IEEE Symposium. 1999. 120 - 132. 10.1109/SECPRI.1999.766909.