IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101226

# Abnormal Event Detection in Videos Using Modified Spatio-Temporal Autoencoder

**Darshan Nemade[1], Paras Thakur[2], Ulkesh More[3], Vaishali Chavhan[4]**

Student, Computer Engineering Department, Shah & Anchor Kutchhi Engineering College, Mumbai, India[1,2,3]

Asst. Professor, Computer Engineering Department, Shah & Anchor Kutchhi Engineering College, Mumbai, India[4]

**Abstract**: The anomaly detection system gives a solution to detect anomaly in crowd event video and sets alarm for public safety in mass gatherings. This paper presents a novel framework to represent video data by a set of general features, which are inferred automatically from a long video footage through a deep learning approach. Specifically, a deep neural network composed of a stack of convolutional autoencoders was used to process video frames in an unsupervised manner that captured spatial structures in the data, which, grouped together, compose the video representation. Then, this representation is fed into a stack of convolutional temporal autoencoders to learn the regular temporal patterns. Our proposed method is domain free (i.e., not related to any specific task, no domain expert required), does not require any additional human effort, and can be easily applied to different scenes. To prove the effectiveness of the proposed method we apply the method to real-world datasets and show that our method consistently outperforms similar methods while maintaining a short running time.

**Keywords**: Anomaly Detection; Convolutional Autoencoders, Deep Learning Technique; Convolutional Neural Network (CNN).

## I. INTRODUCTION

Nowadays the developed countries are improving the security system to defend and manage the public and private crowd. Anomaly detection is a perilous issue in a crowded place. Since Anomaly has made injuries and damages in public area. Sometimes if any anomaly has occurred in a crowded area, the anomaly detection is essential to protect people and the environment without any severe impairment. When the anomaly is perceived, alerting crowd people by an alerting system is very imperative. The alerting system is in different forms such as tones, voice and alert message. After detecting anomaly in the crowd, the alarm system should intimate message or make sound automatically. Particularly in private and public crowded area, the government needs a solution to provide safety now- a- days with low cost. People need security in mass assemblies, public and private events. Thus the Deep Learning based computer vision technique [1], provides a lot of talented methods for private and public safety. And also the technique affords real time video surveillance system for crowd management.

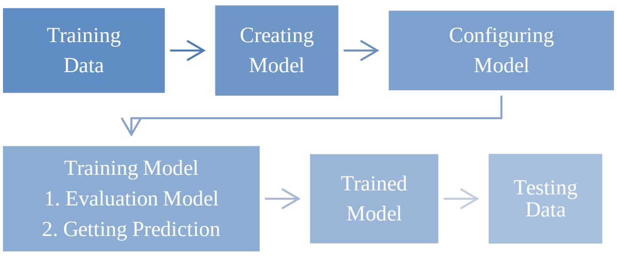


Figure 1 Steps convoluted in the anomaly detection system.

Anomaly detection implements as an essential and requisite phase in the process of assessing the video events (Musical Function, Public Meeting, Bazaar, and Protest). The anomaly detection system will be very cooperative if the enormous detection in event videos on the web can be routinely categorized into predefined classes. Video event holds visual information of anomaly can be detected on a frame basis using Convolution Neural Network (CNN). System has initialized CNN model and implemented with high resolution video event frames. The huge amount of trained data used for working out the CNN mode [2].

The rest of the paper is organized as follows: Section 2 discusses a more brief survey of related works. Section 3 discusses about proposed method and different parameters ueasd to implement proposed model used for anomaly detection. The implementation method and results and extensive experimental results analysis are discussed in Section 4, where we also briefly introduce the new dataset. Finally, Section 5 concludes the paper.

  

IJARCCE
ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101226

## II. LITERATURE SURVEY

### *A. Anomaly Detection Approaches In Video Surveillance*

Various approaches to detect anomalies in surveillance video have been proposed. The choice of a suitable approach is dependent on the nature of data available and also on the environmental characteristics where the surveillance application is deployed. In applications where there is a known behavioral pattern, anomaly detection becomes easier and can be implemented using some rule based approaches where a set of pre-defined rules are coined and fed to the surveillance system. Any deviation from those rules is categorized as anomaly. A rule based anomaly detection approach to detect the anomalies in ship was proposed by (Liu et al 2015). Various parameters such as longitude, latitude, speed and direction were considered to frame the rules that determine the trajectory of movement of ship. An optimal decision rule based approach was proposed by (Saligrama et al 2012) to determine local anomalies and a probabilistic framework was developed.

When the common behavioural pattern is unknown, training based approaches are preferred for detecting anomalies. Training based approaches involve the usage of some set of data to train the system to understand the common behavioural pattern and there by classify any abnormal activities as anomalies. Standard classifier based approaches such as Random Forest, SVM and other classification mechanisms are used to classify anomalies. When the training data is not balanced an ensemble of classifiers are deployed to balance the training data. When the data is auto-correlated, time series based approaches or Recurrent Neural Network based approaches are used. However, the training data may not be available at all times. In such cases, anomaly detection can be accomplished using semi-supervised or unsupervised learning. It may be applying some point based anomaly approaches such as percentiles and histograms or applying some collective anomaly approaches. If the data is univariate in nature, Markov chain based approach or any model based approaches can be deployed to detect anomalies. When the data is multivariate and ordered, a combination of clustering and Markov chain based approaches can be used. If the data is multivariate and un-ordered, any of the clustering based approaches or K-nearest neighbour based approaches can be used. Figure 6 shows the taxonomy of different approaches to anomaly detection.

This literature review was performed over several works related to detecting anomalies such as detecting masked or partially occluded faces, anomaly detection in video sequences, detecting anomalies in crowded area and detecting abandoned objects in video.

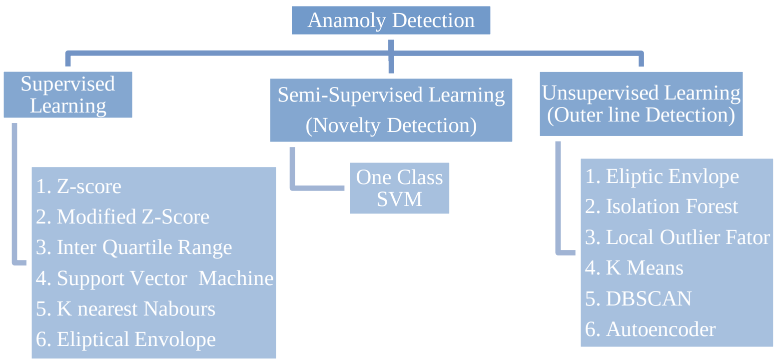


Figure 2 Different approaches used for anomaly detection

### *B. Review on Detecting Anomalies in Video Sequences*

Kim & Reddy (2006) had proposed a network based measurement approach which can spontaneously identify and detect attacks and anomalous traffic by monitoring packet headers passively. Saligrama *et al.*(2010) proposed a family of unsupervised approaches to anomaly detection in videos based on statistical activity analysis.(Li *et al.* 2012) have addressed the automatic anomaly detection problem for surveillance applications by devising a general framework for anomalous event detection in un-crowded sequences. Tran *et al.*(2014) proposed a solution to search for spatio-temporal paths for detecting events in video which can detect and locate video events accurately in cluttered space and at the same time produced stable results to camera motions.

Hu *et al.*(2018) proposed a modified LBP called as squirrel cage LBP (SCLBP) that can encode the motion information effectively and was robust to noise and unwanted disturbances caused by dynamic background and lighting changes. Piciarelli *et al.*(2008) proposed an approach based on single-class Support Vector Machine (SVM) clustering, where the SVM classifier was used for the identification and detection of anomalous trajectories. Piciarelli & Foresti (2011) have worked towards semantically interpreting video sequences to detect anomalous, dangerous or forbidden

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101226

situations. Leyva *et al.*(2017) proposed an approach that used a compact set of highly descriptive features, which was extracted from a new cell structure which helped to define supportive regions from coarse to fine fashion.

Sabokrou *et al.*(2016) introduced two novel cubic patch based anomaly detector approaches where one worked based on power of an auto encoder on reconfiguring an input video patch and another one was based on the sparse representation of an input video patch. Using this, a fast and precise video localisation and anomaly detection method was presented. Laxhammar & Falkman (2014) proposed a sequential Hausdorff Nearest Neighbor Conformal Anomaly Detector (SHNN-CAD) for online learning and sequential anomaly detection in trajectories. This algorithm was having less input parameters and offered a well formed approach to calibrate the anomaly threshold. Mo *et al.*(2014) developed a new joint model based on sparse representation for anomaly detection that enabled the joint anomalies detection involving more than one objects. A greedy pursuit technique was deployed to solve the continuous sparsity problem.

Xiang & Gong (2008) proposed a new framework for automatic behaviour profiling and online detection of anomalies without any manual labelling of the training data set with the aim to address the modelling video behaviour problem captured in surveillance videos for the application of anomaly detection and online normal behaviour recognition. Thomaz *et al.*(2018) developed a family of algorithms based on sparse decompositions that detect anomalies in video sequences obtained from slow moving cameras to restrict search space to the most relevant subspaces search spaces. Cheng *et al.*(2015) presented a hierarchical framework for detecting local and global anomalies via hierarchical feature representation and Gaussian process regression (GPR) which was fully non-parametric and robust to the noisy training data, and supported sparse features.Hu *et al.*(2016) proposed a deep incremental slow feature analysis (D-IncSFA) network which was constructed and applied to directly learning progressively abstract and global high-level representations from raw data sequence. The D-IncSFA network had the functionalities of both feature extractor and anomaly detector that make AD completion in one step.

Ying Zhang *et al.*(2016) proposed a novel anomaly detection approach based on Locality Sensitive Hashing Filters (LSHF), which hashed normal activities into multiple feature buckets with Locality Sensitive Hashing (LSH) functions to filter out abnormal activities. (Emmanu Varghese *et al.*) proposed a new supervised algorithm for detecting abnormal events in confined areas like ATM room, server room etc. (Siqi Wang *et al.* 2018) proposed a novel approach to detect and localize video anomalies automatically. Video volumes were jointly represented by two novel local motion based video descriptors, SL-HOF and ULGP-OF. Sovan Biswas & Venkatesh Babu(2017) proposed a novel idea of detecting anomalies in a video, based on short history of a region in motion based on trajectories. Maying Shen *et al.*(2018) proposed a Nearest Neighbour (NN) based search with the Locality-Sensitive B-tree (LSB-tree) to detect anomalies, which helped to find the approximate NNs among the normal feature samples for each test sample. Dan Xu *et al.*(2014) proposed an approach to detect anomalies based on a hierarchical activity pattern discovery framework, comprehensively considering both global and local spatio-temporal contexts. Tian Wang *et al.*(2018) proposed an algorithm to solve abandoned object detection efficiently based on an image descriptor which encodes the movement information and the classification method.

Huorong Ren *et al.*(2017)proposed an anomaly detection approach based on a dynamic Markov model. This approach segmented sequence data by a sliding window. Also, an anomaly substitution strategy was proposed to prevent the detected anomalies from impacting the building of the models and keep anomaly detection continuously. Fan Jiang *et al.*(2011) proposed a hierarchical data mining approach where frequency-based analysis was performed at each level to automatically discover regular rules of normal events. Events deviating from these rules were identified as anomalies. Shifu Zhou *et al.*(2016)coupled anomaly detection with a spatial–temporal Convolutional Neural Networks (CNN) to capture features from both spatial and temporal dimensions by performing spatial–temporal convolutions, thereby, both the appearance and motion information encoded in continuous frames were extracted.

## III. EXPERIMENT AND RESULTS.

This project is differs from other implementations in quite a few aspects. Firstly you can input any video based data for training. The data will be pre processed and converted into numpy matrix which will be used for training. Converting data into matrix allows for more efficient retention of data. We have also tinkered with the layers and parameters of the model to make it more efficient and accurate.

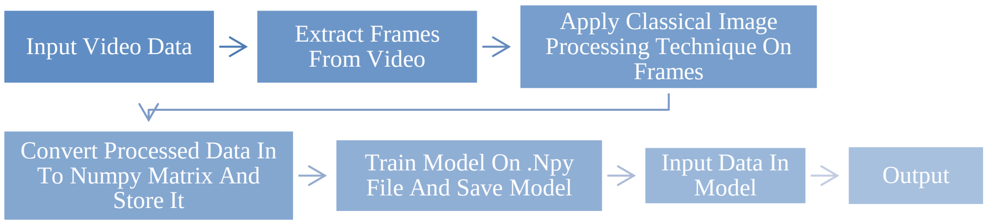


Fig. 3. Algorithm of proposed system

Our changes led to a boost in accuracy compared to baseline. We have also added 4 different ways to deploy the model. A person can use frames extracted from video to get accuracy, use real time video feed, use saved video or directly use

  

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101226

a numpy file of the data to be classified and pass it through the model. All these changes make our project flexible, adaptable and modular.

*A. Dataset*

Avenue Dataset contains 16 training and 21 testing video clips. The videos are captured in CUHK campus avenue with 30652 (15328 training, 15324 testing) frames in total. This dataset contains the some challenges. Slight camera shake (in testing video 2, frame 1051 - 1100) presents. A few outliers are included in training data and Some normal patterns seldom appear in training data.

*B. Implementation work:*

The famous recent technique of Modified Spatio-Temporal Autoencoder technique is applied in the proposed with help softwares such as Visual Studio 2012 Various Dependencies and library based Dependencies such as ffmpeg for Video frame extraction, numpy, sklearn, keras, tensorflow, h5py, scipy, OpenCV and hardware used Computer with windows OS: for simulation and training and code compilation purpose, Camera: record real time anomaly activity performed by subject and Pen drives: transfer data from one device to another device with python code.

1. Parameters:

In the proposed system, the convolutional network model is constructed with some crucial parameters. The three convolution layers are implemented by the activation function of layers namely, Relu and max pooling layer. In this layer, it has filters. The kernel size is 2×2. The model is trained for four classes. There are four neurons in the output layers. The special activation function of this network for classifying the dataset is categorical-cross entropy.

2. Training:

The training phase of this work has 6 epochs and 100 training samples to implement the model for extracting crucial features and good training. The needed dataset of video frames are stored in a stack array and modified size as 150 x150. The stacked array of the dataset is changed into a batch file and provides data to the Spatio-Temporal Autoencoder model for the training process. The model extracts the features through epochs to detect the anomaly using separate labels (0,1,2,3) for anomalies namely.

3. Testing:

The final phase of testing in the Spatio-Temporal Autoencoder model detecting the anomaly is in different video events is taken and converted these into frames. 100 datasets of video events are stored in a stack array of a batch file and modified size as 150x150. The event video frames are collected from different events namely Sports, Protest, Temple, etc. From each video, 30 frames are collected and stored for a test container. In that, 10 false datasets are collected from other videos and stored in the test container. From the batch file, the testing data is sent to the trained model. The Spatio-Temporal Autoencoder model finds four categories of the anomaly and shows anomaly name for each category correctly. Shown in Figure 5. The Spatio-Temporal Autoencoder baseline gives 100% validation accuracy

The main goal of the proposed system is used to acquire in what way to recognize anomaly on various crowd videos. The proposed system has applied Spatio-Temporal Autoencoder baseline and VGG-16 for crowd video anomaly detection. The network model predicts the scenes of the test images at that time. The performance of the Spatio-Temporal Autoencoder baseline and VGG-16 model is evaluated and calculated as true positive, true negative, false positive and false negative in Table 1. Totally 100 true datasets and 30 false datasets are taken to classify the anomaly from crowd video.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101226


Fig. 4. Frame Sequence of a segmented video


Fig. 5. Man throwing bag in air


Fig. 6. Small Boy Jumping


Fig. 7. Man Running Man


Fig. 8. Running In Opposite Direction

## IV. RESULT ANALYSIS

### *A. Performance Parameters*

The performance of the existing and the proposed methods for human activity prediction in the VSS are analysed on the basis of accuracy, precision, recall, information gain ratio, and true positive rate.

#### a) True Positive Rate

True Positive Rate is described as the rate of human abnormal activity predicted as human abnormal activity in videos.

#### b) Accuracy

It is the fraction of true results of human activity prediction (true positive and true negative) among the total number of cases analyzed. It is calculated as,

$$Accuracy = \frac{True\ Positive\ (TP) + True\ Negetive\ (TN)}{TP + TN + False\ Positive\ (FP) + False\ Negetive\ (FN)}$$

**Equation** Error! No text of specified style in document..**1**

where, if the class label is positive and the human abnormal activity prediction outcome is positive, then it is TP. If the class label is negative and the human abnormal activity prediction outcome is negative, then it is TN. If the class label is negative and the human abnormal activity prediction outcome is positive, then it is FP. If the class label is positive and the human abnormal activity prediction outcome is negative, then it is FN.

#### c) Precision

It is the fraction of the number of suspicious faces that are appropriately recognized to the sum of the count of correctly recognized suspicious faces and the wrongly recognized suspicious faces.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 12, December 2021
DOI: 10.17148/IJARCCE.2021.101226

$$Precision = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Positive\ (FP)}$$ **Equation** Error! No text of specified style in document..**2**

**d) Recall**

It is the fraction of the number of suspicious faces that are appropriately recognized to the sum of the count of correctly recognized suspicious faces and the wrongly recognized non-suspicious faces.

$$Recall = \frac{True\ Positive\ (TP)}{True\ Positive\ (TP) + False\ Negetive\ (FN)}$$ **Equation** Error! No text of specified style in document..**3**

**e) Information Gain Ratio**

It is defined as a quantity of knowledge obtained during the prediction of human activities in the videos.

**f) Regularity Score**

Once the model is trained, we can evaluate our models performance by feeding in testing data and check whether it is capable of detecting abnormal events while keeping false alarm rate low. To better compare with [5], we used the same formula to calculate the regularity score for all frames, the only difference being the learned model is of a different kind. The reconstruction error of all pixel values I in frame t of the video sequence is taken as the Euclidean distance between the input frame and the reconstructed frame:

$$e(t)\|\mathbf{x}(\mathbf{t}) - \mathbf{fW}(\mathbf{x}(\mathbf{t}))\|\mathbf{2}$$ **Equation** Error! No text of specified style in document..**4**

where fW is the learned weights by the spatiotemporal model. We then compute the abnormality score sa(t) by scaling between 0 and 1. Subsequently, regularity score sr(t) can be simply derived by subtracting abnormality score from 1:

$$sa(t) = \frac{e(t) - e(t)min}{e(t)max}$$ **Equation** Error! No text of specified style in document..**5**

$$sr(t) = 1 - sa(t)$$ **Equation** Error! No text of specified style in document..**6**

The unusual circumstances comprise of various volunteers suddenly dancing, running and pushing in a crowded place. Overall there are six kinds of unusual or abnormal circumstances which take place in 12000 frames of video series. The usual screening quality for a video surveillance is 720 576 by 29 frames per second, which is the spatial motion of a novel video frame. Moreover, a different strategy which is projected by the authors in (Jian-hao&Li 2011) is contrasted with the presentation of this strategy and the resulting outcome is evaluated. Each frame is divided into four segments in our present research work. The number of segments per frame is customizable. The entropy of DCT coefficients is computed for every segment also the median rate for the first 500 frames is calculated. In relation to this research and analysis, the threshold median entropy is positioned to 3 times than the median rate to categorize the abnormal happenings. If there are any unusual happenings in any of the segment, in such cases an unusual indicator raises for the entire structure. Table 1 represents the set of all frames extracted from a segmented video. The duration of the segmented video is one minute and is customizable and obtained results are summarised in table I below figure 9 shows comparisons of parameters summarised in table below parameters such as Accuracy, precision and recall as per results accuracy of all sample video analysed is 50% and Precision is 53% and recall value is 0.8 approx.

TABLE I
SYSTEM ACCURACY FOR DIFFERENT DATA SAMPLES TAKEN

| Data (video) | Frames Classification | | | | Total frames | Accuracy | Precision | Recall |
|---|---|---|---|---|---|---|---|---|
| | True Positive | False Positive | True Negative | False Negative | | | | |
| **Boy Jumping** | 435 | 430 | 65 | 70 | 500 | 0.5035 | 0.5028 | 0.8613 |
| **Man Running** | 16 | 10 | 314 | 330 | 330 | 0.4925 | 0.6153 | 0.0462 |
| **Man Running Opposite** | 634 | 630 | 136 | 140 | 770 | 0.5024 | 0.5015 | 0.8191 |
| **Throwing Bag** | 563 | 560 | 537 | 540 | 1100 | 0.5045 | 0.5013 | 0.5104 |

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101226

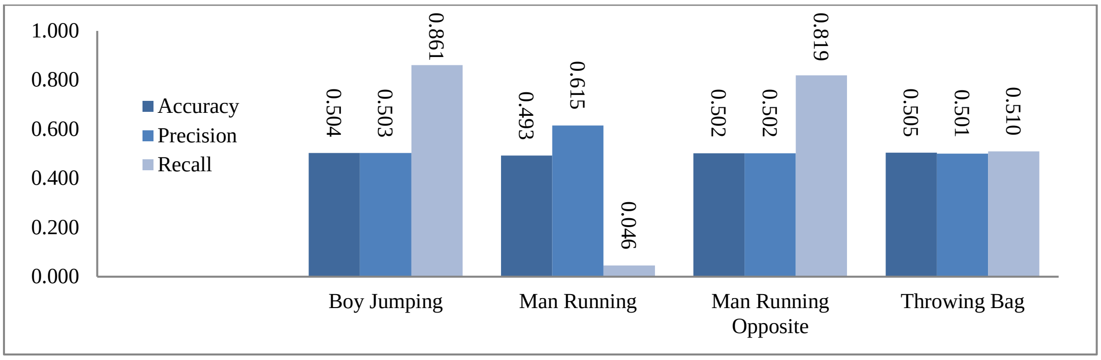


Fig. 9. Comparison graph of Accuracy, Precision and Recall

## V. CONCLUSION

The proposed system detects anomaly using Spatio-Temporal Autoencoder based techniques. After detecting the anomaly region, automatically the system makes visual alert in the form of massege and alarm with frame detection. The technique works well and finds the different anomalies such as WalkFall, ClimbLadder, JumpOverGap, PullHeavyObject, Kick, ShotGunCollapse, LookInCar, PickupThrowObject, WalkTurnBack, DrunkWalk, CrawlOnKnees, WaveArms, DrawGraffiti, JumpOverFence, RunStop, SmashObject, and Punch in different location. The proposed system can implement in large space crowd area and parameters such as Accuracy, precision and recall as per results accuracy of all sample video analysed is 50% and Precision is 53% and recall value is 0.8 approx. In future work, the system could be considered for implementing more different types of anomaly.

## ACKNOWLEDGMENT

We wish to express our profound gratitude to our principal Dr. Bhavesh Patel for allowing us to go ahead with this project and giving us the opportunity to explore this domain. We would also like to thank our Head of Department Prof. Uday Bhave for our constant encouragement and support towards achieving this goal. We would also like to thank the Review Committee for their invaluable suggestions and feedback without whom our work would have been very difficult. We take this opportunity to express our profound gratitude and deep regards to our guide Mrs. Bhakti Sonawane for her exemplary guidance, monitoring and constant encouragement throughout the course of this project. The blessing, help and guidance given by her time to time shall carry us a long way in the journey of life on which we are about to embark. No project is ever complete without the guidelines of these experts who have already established a mark on this path before and have become masters of it. So, we would like to take this opportunity to thank all those who have helped us in implementing this project.

## REFERENCES

[1] Adam, A., Rivlin, E., Shimshoni, I., Reinitz, D.: Robust real-time unusual event detection using multiple fixed-location monitors. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(3), 555–560 (2008)

[2] Cong, Y., Yuan, J., Liu, J.: Sparse reconstruction cost for abnormal event detection. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 3449–3456 (2011)

[3] Doll´ar, P., Rabaud, V., Cottrell, G., Belongie, S.: Behavior recognition via sparse spatio-temporal features. In: Proceedings - 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, VS-PETS. vol. 2005, pp. 65–72 (2005)

[4] Hasan, M., Choi, J., Neumann, J., Roy-Chowdhury, A.K., Davis, L.S.: Learning temporal regularity in video sequences. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 733–742 (June 2016)

[5] Ji, S., Yang, M., Yu, K.: 3D Convolutional Neural Networks for Human Action Recognition. Pami 35(1), 221–31 (2013), http://www.ncbi.nlm.nih.gov/pubmed/22392705.

[6] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L.: Large-scale video classification with convolutional neural networks. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. pp. 1725–1732 (June 2014)

[7] Kim, J., Grauman, K.: Observe locally, infer globally: A space-time MRF for detecting abnormal activities with incremental updates. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009. pp. 2921–2928 (2009)

[8] Kozlov, Y., Weinkauf, T.: Persistence1d: Extracting and filtering minima and maxima of 1d functions. http://people.mpi-inf.mpg.de/~weinkauf/notes/persistence1d.html, accessed: 2017-01-05.

[9] Kratz, L., Nishino, K.: Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops,CVPR Workshops 2009. pp. 1446–1453 (2009)

[10] Laptev, I., Marsza lek, M., Schmid, C., Rozenfeld, B.: Learning realistic human actions from movies. In: 26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR (2008)

[11] Li, C., Han, Z., Ye, Q., Jiao, J.: Abnormal behavior detection via sparse reconstruction analysis of trajectory. In: Proceedings - 6th International Conference on Image and Graphics, ICIG 2011. pp. 807–810 (2011)

[12] Lu, C., Shi, J., Jia, J.: Abnormal event detection at 150 fps in matlab. In:2013 IEEE International Conference on Computer Vision. pp. 2720–2727 (Dec 2013)

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 12, December 2021

DOI: 10.17148/IJARCCE.2021.101226

[13] Mahadevan, V., Li, W., Bhalodia, V., Vasconcelos, N.: Anomaly detection in crowded scenes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp. 1975–1981 (2010)
[14] Medel, J.R.: Anomaly Detection Using Predictive Convolutional Long Short-Term Memory Units. Master's thesis, Rochester Institute of Technology (2016), accessed from http://scholarworks.rit.edu/theses/9319.
[15] Mehran, R., Oyama, A., Shah, M.: Abnormal crowd behavior detection using social force model. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009. pp. 935–942 (2009) Mo, X., Monga, V., Bala, R., Fan, Z.: Adaptive sparse representations for video anomaly detection. IEEE Transactions on Circuits and Systems for Video Technology 24(4), 631–645 (2014)
[16] Oneata, D., Verbeek, J., Schmid, C.: Action and Event Recognition with Fisher Vectors on a Compact Feature Set. 2013 IEEE International Conference on Computer Vision pp. 1817–1824 (2013), http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6751336.
[17] Patraucean, V., Handa, A., Cipolla, R.: Spatio-temporal video autoencoder with differentiable memory. International Conference On Learning Representations (2015), 1–10 (2016), http://arxiv.org/abs/1511.06309.
[18] Piciarelli, C., Micheloni, C., Foresti, G.L.: Trajectory-based anomalous event detection. IEEE Transactions on Circuits and Systems for Video Technology 18(11), 1544–1554 (2008)
[19] Reddy, V., Sanderson, C., Lovell, B.C.: Improved anomaly detection in crowded scenes via cell-based analysis of foreground speed, size and texture. In: CVPR 2011 WORKSHOPS. pp. 55–61 (June 2011)
[20] Roshtkhari, M.J., Levine, M.D.: An on-line, real-time learning method for detecting anomalies in videos using spatio-temporal compositions. Computer Vision and Image Understanding 117(10), 1436–1452 (2013)