

Video Anomaly Detection using Likelihood Statistical Texture Feature Representation in Surveillance Video

Sujith Kumar P. S¹, Manish T I², Vince Paul³ and Sanaj M S⁴

¹Professor & Dean, Dept of CSE, Sree Buddha College of Engineering, Pattoor – 690529

²Professor & Head, Dept of CSE, SCMS School of Engineering and Technology – 683576

³Professor & Head, Dept of CSE, Christ College of Engineering, Irinjalakuda- 680125

⁴Associate Professor, Dept of CSE, Adi Shankara Institute of Engg & Technology, Kalady - 683574

*Correspondence: research_phd@myyahoo.com

ABSTRACT- An anomaly detection system is challenging to develop due to the non-deterministic nature and lack of a clear definition of abnormal events. To address this issue, this paper introduces the Likelihood Statistical Texture Feature Representation (LSTFR) method, which uses CSR (Co-occurrence with Stationary occurrences Representation) to construct spatial activity patterns using gray-level co-occurrence matrices with likelihood estimation. Additionally, LSTFR constructs a composition histogram representation to model normal behavior. The occurrence rate in LSTFR is characterized by a histogram representation that depends on the spatio-temporal information of the frame sequence. To efficiently classify events using LSTFR, this paper employs Convolutional Long Short-Term Memory (conv-LSTM), where the histogram representation of LSTFR is automatically modeled through training with normal events. The efficiency of the proposed approach is evaluated using three datasets: UMN, Subway, and Avenue. Finally, the results of the proposed method are compared with several existing algorithms.

Keywords: Abnormal event detection, co-occurrence matrix, likelihood estimation, and conv-LSTM.

ARTICLE INFORMATION

Author(s): Sujith Kumar P. S, Manish T I, Vince Paul and Sanaj M S;

Received: 00/00/2024; **Accepted:** 00/00/2024; **Published:** 30/06/2024;

e-ISSN: XXXX-XXXX;

Paper Id: IJCSR-030106;

Citation: 10.37391/IJCSR.030106



Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

The importance of video surveillance systems has increased significantly recently due to their critical role in enhancing security. These systems are extensively used in various applications such as educational institutions, healthcare facilities, public places, and traffic analysis. In many cases, surveillance systems rely on human observers to monitor video footage. These observers must be vigilant and attentive to the activities within the surveillance area. However, due to limitations in human resources and the need for numerous cameras, the accuracy of security can be affected by the observer's alertness [1]. To enhance the security system's accuracy, there is a growing need for automated surveillance intelligence systems capable of detecting suspicious human behaviors and abnormal events.

Generally, all would be interested to know if any suspicious events occurred in the frame sequence. However, it is highly hard to design and develop abnormal event approaches with pre-defined rules [2]. Jiang et al. [3] developed an abnormal event detection system using motion patterns. They also define abnormal events as rarely occurring in the frame sequence compared to regular behaviours events. However, making these

concepts requires semantic knowledge as well as subjective knowledge about the frame sequence. In most of the surveillance environment, making the abnormal event's pattern would be very hard to represent by mathematical logic. Therefore, abnormal event detection is built using frequent normal patterns that would be found in the frame sequence. However, if any frame is not fit to any normal pattern, the algorithm concludes a given frame sequence's event as an abnormal event [4]. Some of the abnormal events are explained here: 1. Human movement with high speed, 2. Vehicles moving in the pedestrian path, 3. Humans fell down, 4. Human loitering, 5. Humans running in different directions, 6. Many humans merged as a crowd in a local region, 7. Human intrusion. 8. Intruders in Region Of Interest (ROI) or restricted areas. To detect the above abnormal events, it needs an efficient spatio-temporal descriptor. There are several descriptors available in the literature survey such as HGOFO [5], HOFOM [6-8], HOG3D [9], MBH [10], etc. Although these descriptors work in both spatial and temporal content they cannot well describe the object's texture in many situations due to lack of structure of descriptor information.

To achieve the efficient abnormal event detection system, this paper introduced the LSTFR method to describe the structure of the objects in a frame sequence. This method develops the spatio-temporal pattern to recognize the event type whether normal crowd events or abnormal crowd events. The pattern is developed using three sub works which are used as the clue to differentiate normal event patterns and abnormal event patterns in given a scene: i) co-occurrence with stationary-occurrence and likelihood estimation – find a joint probability of each pairs as well as stationary pair, ii) same probability of pairs are computed – finds spatial structure of a frame, and iii) constructs composition information histogram – characterises the spatial

and temporal activity pattern. The contributions of the paper are summarized here:

1. The spatio-temporal pattern, namely LSTFR is developed to derive the spatial and temporal information from the frame sequence to accurately recognize the event types.
2. The conv-LSTM is modelled based on the LSTFR feature vector to efficiently classify the events.
3. In order to find the accurate classification results, the conv-LSTM parameters are tuned to accurately converge the feature vector values for conv-LSTM.

The sections of the paper are classified as follows. In section 2, explains the related work on abnormal event. In section 3, explains the LSTFR methodology. In section 4, discuss the results along with comparative analysis. In section 5, explains the conclusion and future direction.

2. RELATED WORK

This section discusses the procedure on existing abnormal event system. The relevant work is classified into two sub categories viz., trajectory based abnormal event and motion based abnormal events. The trajectory based abnormal event is done by using object tracking analysis. It requires a sparse environment to detect and track the object in the subsequent frames. Motion based abnormal events are suitable for density or crowd surveillance environments. The motion will be analyzed for a crowded frame sequence by analyzing people movement without describing objects individually [4], [8]. The accuracy decreases if the more number of persons appearing in the video in case of trajectory based approach [11-16]. Generally, motion based analysis is applied on the crowd scene that tends to discover crowd's behaviour based on their attributes and size. The trajectory based abnormal event [14-15] requires detection and segmentation of object individuals. The trajectory based approaches in most cases are not suitable since objects with occlusion in high density crowd scenes. The motion based approaches use dense information rather than determining target objects. The motion based approaches mainly focus on the crowd movement and crowd occupies how much spatial location in the 2D dimensional in the vision field.

There are two main common steps such as representation and classification involved in abnormal event detection. The trajectory based abnormal event employs the appearance of object's feature representation such as color chromatic content, edge feature [16], textures analysis [17], contextual spatial information [18], and motion feature based on Optical Flow (OF) estimation [4], [8]. Also, saliency maps such as the lagrangian particle map is extracted for describing crowd representation [19] and [20].

To develop and predict both trajectory and motion based abnormal events, statistical measures viz., Gaussian Mixture Models (GMM), Hidden Markov models (HMM), Expectation Maximization are used to find the unusual events in the certain region [16], [21], [22]. Antic and B. Ommer developed a model [23] using feature extraction and Gaussian Regression Process

[24]. Cheng [25] developed a method for detecting abnormal crowds using a hierarchical mixture textures analysis to represent the nature of the content. The descriptors are developed [10], [26] they share similar clues about the descriptor. Shao et al. [27] developed an algorithm using group profiling by extracting four descriptors to define the abnormal event. Bag of Features (BOF) and SVM are used to quantize the information and classify the given event respectively. Cheng et al. [28] proposed a BOF using codebook representation. In recent days, many studies used a CNN to train the classification model using spatio-temporal patches that is derived from the optical flow information on the frame sequence [29]. Yuan et al. [30] developed a method using SCD to explain the individual of the crowd. This model worked as a multi-object tracker to connect the target object in subsequent frames. The selective histogram of OF is used to find the difference of motion between individuals. The histogram representation is used as the motion statistics which consists of optical flow information. From the literature survey, this paper addresses that the spatio-temporal structure of the content is needed to develop the abnormal event detection methods.

3. THE LSTFR METHODOLOGY

This paper presented a novel LSTFR video anomaly detection method which builds upon likelihood statistical texture estimation which characterizes the spatial activity pattern in a frame sequence. The LSTFR is derived from co-occurring and stationary-occurring grey level intensities. The LSTFR is used for evaluating the neighbourhood information that evolves over time period. In order to provide an efficient feature descriptor of crowd environments, the LSTFR is developed.

3.1 Problem Formulation

Let V be the surveillance video which consists of a frame sequence $H = \{h_1, h_2, h_3, \dots, h_N\}$. To efficiently detect the content of the frame sequence from light scattering or reflection, H is needed to be converted into grey scale format using [31] the equation 1.

where R, G, B represents the color pixels IR, IG, and IB that are the red, green and blue of RGB respectively. The output of the gray scale image can be represented as $G = \{g_1, g_2, g_3, \dots, g_n\}$.

An inefficient spatial representation cannot well work to characterize the nature of the video. Therefore, analyzing pixel occurrence is one of important methods to characterize the spatial activity pattern. To analyze the spatial activity pattern in the local region, this paper used region based analysis [32]. Therefore, the total frames N are to be divided into $K \times L$ with non overlapping region W_i with similar size where $i = 1$ to n . The region is divided a smaller 5×5 -pixel local image area within the frame to efficiently analyze the content. Then, Co-occurrence with Stationary-occurrence Representation (CSR) is derived from the Goneach region W_i to estimate spatial activity patterns.

3.1.1. Spatial Activity Construction

The CSR is a feature representation and it is built by counting and analysing the occurrence of pixel pairs between current and

the consecutive pixels [33]. Also, to efficiently compute the occurrence, stationary-occurrence is also need. Stationary-occurrence is computed by counting the end row pixel pair occurrence, in which listed out that occurrences and combined with the respective neighbourhood occurrence.

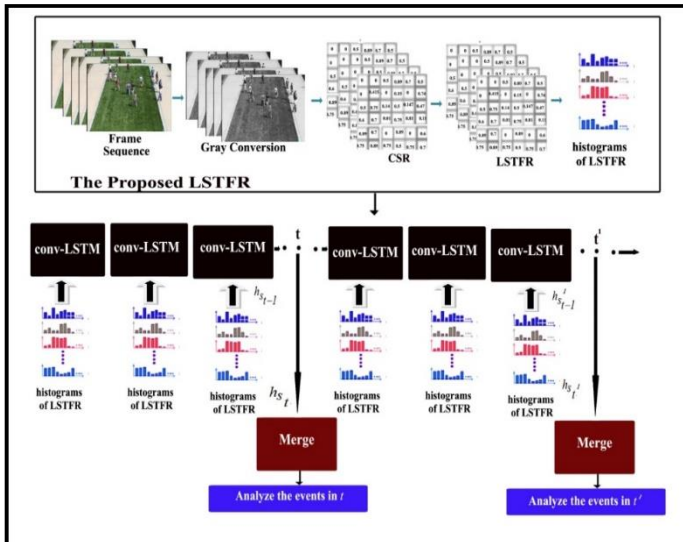


Figure. 1 The proposed LSTFR workflow

This process can efficiently describe the texture on each local region. The relationship of pixel pair is determined by using the displacement vector d and direction k . The d and k are applied on the pixel pairs of current and the consecutive pixels. In CSR, the value selection for both d and k is important since these values can reflect the realistic content by CSR. Therefore, this work fixed the minimum value for both as ($d = 0^\circ, k = 1$). Because of minimum d and direction k can analyze efficient occurrences on the G . The output CSR needs to be normalized for holding the statistical information. The resultant CSR is represented in equation (2).

$$CSR = \begin{bmatrix} I_{11} & I_{12} & I_{13} & \dots & \dots & I_{1n} \\ I_{21} & \dots & \dots & \dots & \dots & I_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ I_{n1} & \dots & \dots & \dots & \dots & I_n \end{bmatrix} \quad (2)$$

where CSR is a square matrix and $I_{11}, I_{12}, \dots, I_n$ are the degrees of occurrence of current and consecutive pair pixels.

3.2 Likelihood Statistical Texture Features Representation

We compute the LSTFR features from the CSR on each W_i by the likelihood estimation. Likelihood estimation is called a joint probability that calculates the likelihood of any two pairs in CSR occurring together. It is computed by the equation (3).

$$\frac{P((I_{a1}, I_{b1}) \cap (I_{a2}, I_{b2}))}{P(I_{a2}, I_{b2})} \cdot \frac{P((I_{a1}, I_{b1}) \cap (I_{a3}, I_{b3}))}{P(I_{a3}, I_{b3})} \dots \frac{P((I_{a_{n-1}}, I_{b_{n-1}}) \cap (I_{an}, I_{bn}))}{P(I_{an}, I_{bn})} \quad (3)$$

where (I_{a1}, I_{b1}) and (I_{a2}, I_{b2}) are the pairs for analysing joint probability.

Let L_{mn} be the joint probability values of each pair W_i where m and n are the row and column vectors respectively. Using these probability values, the Similar Joint Probability (SJP) value is computed for each two states to characterize the spatial information. The output L_{mn} can be described by following equation (4).

$$L_{mn} = \begin{bmatrix} l_{11} & l_{12} & \dots & \dots & l_{1m} \\ l_{21} & l_{22} & \dots & \dots & l_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ l_{nm} & l_{nm} & \dots & \dots & l_{nm} \end{bmatrix} \quad (4)$$

To estimate the SJP, the distance (ds) measure is applied on the L_{mn} .

3.3 Constructing the Composition Histogram

From the above LSTFR, we can compute a histogram representation by counting the occurrence probability values of each W_i within a frame. This histogram features the ability to characterise the spatial activity patterns. However, this type of histogram feature expresses different size feature vectors and it leads to difficult comparison between any two histogram vectors due to there is no systematic arrangement about the SJP. To overcome this challenge, this work constructs the composition histogram representation by quantizing into 24 bins with equal intervals. Binning is a process [34] of counting each SJP values into a number of bin ranges. The bin ranges are arranged in the same position of each histogram h_s .

3.4 Abnormal Crowd Events Detection

To detect normal or abnormal events, high-dimensional features are needed. In this context, the abnormal events of the frame are analyzed using conv-RNN. The conv-RNN is a kind of neural network which is used for modelling the feature vector. In general, conv-RNN permits the connections with hidden units that can be associated with time delay. This process is useful to encode the historical feature information and compute the relationship between features. However, a traditional conv-RNN model is difficult to train the frame sequence since this model cannot manage correctly long-term dynamics due to declining Stochastic Gradient Descent (SGD). The SGD is used to find the model parameters which correspond to the best fit between expected and actual output. In order to handle the memory problem, this paper modelled conv-LSTM[35] based on the LSTFR. The proposed LSTFR is explained in figure 1 where current and previous time t and $t-1$ respectively and current and previous cell status are h_{s_t} and $h_{s_{t-1}}$ respectively.

4. EXPERIENTIAL SETUP AND RESULTS

The LSTFR is experimented using MATLAB 13a software to evaluate the results by comparing the several existing abnormal event detection approaches. The three challenging datasets such as UMN [36], Subway [37], and Avenue are used to classify the events type using LSTFR with conv-LSTM model. To train the conv-LSTM model, the LSTFR feature vector is applied with the batch size of 128, epochs of 200, and stride 2.

4.1 Performance metric

To evaluate the performance, this paper used Receiver Operating Characteristic (ROC). To summarize the performance results, using the corresponding Area Under Curve AUC. Further, the Equal Error Rate (EER) is used to estimate the equal probability for misclassifying samples and the misclassification may be both negative and positive samples in the ROC curve. The above all metrics are applied on the frame level criterion to evaluate anomaly scores of each frame.

4.2 Result Analysis

In the UMN Web abnormality dataset, normal and abnormal frequently occur in indoor and outdoor. The proposed approach detected abnormal frames in appropriate places. Table 1 shows the performance analysis of LSTFR on UMN dataset with several existing abnormal activities approaches such as optical flow [38], SF [39], MDT [40], Chaotic invariants [41] Biswas [42], GLCM [43], OPLKT-EMEHO [44], and rpNet [45].

Table 1 Comparative Analysis on UMN

Methods	Year	AUC
SF	2009	93.00%
MDT	2010	95.00%
Chaotic invariants	2010	96.04%
Optical Flow	2011	86.00%
Biswas	2016	96.00%
GLCM	2017	96.09%
OPLKT-EMEHO	2023	97.00%
rpNet	2023	96.00%
LSTFR		98.00%

The experimental results are carried out on the subway dataset using AUC and EER evaluation. The proposed LSTFR approach is compared with several existing algorithms. The proposed approach obtains better results for both entrance and exit gates when compared with all existing algorithms. Table 2 shows performance results on the subway dataset.

Table 2 Comparative Analysis on Subway

Methods	Year	Subway (Entrance)		Subway (Exit)	
		AUC	EER	AUC	EER
[39]	2009	66.4%	30.2%	56.4%	41.02%
[46]	2013	82.01%	21.4%	83.5%	16.06%
[47]	2016	-	-	86.08%	6.06%
[48]	2016	93.02%	25.09%	81.07%	9.01%
[49]	2017	69.05%	-	83.06%	-

[50]	2020	90.20%	22.67%	94.6%	9.30%
[51]	2023	90.80%	23.50%	93.00%	10.05%
[52]	2023	89.20%	21.60%	94.00%	8.90%
LSTFR		94.80%	20.54%	97.40%	7.60%

The quantitative analysis is done between the LSTFR approach and several existing algorithms using AUC and EER. Many research works have taken the avenue dataset to detect the abnormal event because it includes many challenging environments. The proposed LSTFR yields better results on the Avenue dataset when compared with existing algorithms. Table 3 shows performance results on the Avenue dataset.

Table 3 Comparative Analysis on Avenue Dataset

Methods	AUC	EER
[7]	86.04%	21.08%
[53]	89.2%	17.6%
[51]	90.4%	-
[52]	90.60	-
LSTFR	91.1%	17.4%

4.3 Results and Discussions

The LSTFR approach demonstrates the results better than the existing methods results, which are explained in the table 1, table 2, and table 3 by using three datasets namely UMN, Subway, and Avenue. The detected abnormal events such as humans running in the same time in UMN dataset and human walking in wrong direction and no payment in the Subway dataset. The selection values for the parameters d, k, and t are well apt for all datasets. The parameters are fixed after validating the experimental results.

5. CONCLUSION

This work presented a novel LSTFR method that can determine normal activity patterns to efficiently detect abnormal events in surveillance videos. The LSTFR method constructs spatial activity pattern descriptor using gray level co-occurrence matrix with likelihood estimation. Also, the composition histogram representation is derived from the likelihood estimation to learn the LSTFR feature pattern using the conv-LSTM deep neural network. Also, the LSTFR feature vector with conv-LSTM model revealed that the higher detection accuracy in three datasets namely UMN, Subway, and Avenue. The proposed LSTFR approach efficiently worked in abrupt activities while it does not efficiently work in gradual abnormal activities since the spatial activity pattern descriptor is not enough to express the less abnormality score. The LSTFR proposed approach can be used to detect crowd activities, collision detection, etc.

REFERENCES

- [1] Keval H 2006 Cctv control room collaboration and communication: Does it work?. Human Centred Technology Workshop.
- [2] Aggarwal J and Ryoo M 2011 Human activity analysis: A review, ACM Comput. Surv.
- [3] Jiang F, Wu Y and Katsaggelos A 2009 Detecting contextual anomalies of crowd motion in surveillance video, Image Processing (ICIP). 2009 16th IEEE International Conference.

- [4] Ryan D, Denman S, Fookes C and Sridharan S 2011 Textures of optical flow for real-time anomaly detection in crowds, *Advanced Video and Signal-Based Surveillance (AVSS)*, 2011 8th IEEE International Conference.
- [5] Chaudhry R, Ravichandran A, Hager G and Vidal R 2009 Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions, *Computer Vision and Pattern Recognition, CVPR 2009. IEEE Conference*.
- [6] Colque R V H M C, Junior A C and Schwartz W R 2015 Histograms of optical flow orientation and magnitude to detect anomalous events in videos, *28th SIBGRAPI Conference on Graphics, Patterns and Images, SIBGRAPI 2015, Salvador, Bahia, Brazil*.
- [7] W. Liu, W. Luo, D. Lian, and S. Gao, "Future frame prediction for anomaly detection - A new baseline," in *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, 2018*.
- [8] Adam, Rivlin E, Shimshoni I and Reinitz D 2008 Robust real-time unusual event detection using multiple fixed-location monitors, *Pattern Analysis and Machine Intelligence, IEEE Transactions*.
- [9] Kl'aser, Marszalek M and Schmid C 2008 A spatio-temporal descriptor based on 3d-gradients, *British Machine Vision Conference*.
- [10] Wang H, Klser A, Schmid C and Liu C.-L 2013 Dense trajectories and motion boundary descriptors for action recognition, *International Journal of Computer Vision*.
- [11] Popoola O and Wang K 2012 Video-based abnormal human behaviour recognition; a review, *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions*.
- [12] Vishwakarma S and Agrawal A 2013 A survey on activity recognition and behavior understanding in video surveillance, *The Visual Computer*.
- [13] Choi W and Savarese S 2012 A unified framework for multi-target tracking and collective activity recognition, in *Proceedings of the 12th European Conference on Computer Vision - Volume Part IV*.
- [14] Shao J, Loy C. C and Wang X 2014 Scene-independent group profiling in crowd, in *Computer Vision and Pattern Recognition (CVPR)*, 2014 IEEE Conference.
- [15] Yuan Y, Wang D and Wang Q 2017 Anomaly Detection in Traffic Scenes via Spatial-Aware Motion Reconstruction, *IEEE Transactions on Intelligent Transportation Systems*, pp. 1198-1209.
- [16] Kratz L and Nishino K 2009 Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models, in *Computer Vision and Pattern Recognition, CVPR 2009. IEEE Conference*.
- [17] Guogang Xiong, Jun Cheng, Xinyu Wu, Yen-Lun Chen, Yongsheng Ou and angsheng Xu 2012 An energy model approach to people counting for abnormal crowd behavior detection, *Elsevier, Neurocomputing*, pp. 121 – 135.
- [18] Cong Y, Yuan J and Liu J 2011 Sparse reconstruction cost for abnormal event detection, *Computer Vision and Pattern Recognition (CVPR)*, 2011 IEEE Conference.
- [19] Sun X, Yao H, Ji R, Liu X and Xu P, 2011 Unsupervised fast anomaly detection in crowds, in *Proceedings of the 19th ACM International Conference on Multimedia*.
- [20] Ali S and Shah M 2007 A lagrangian particle dynamics approach for crowd flow segmentation and stability analysis, *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference*.
- [21] Andrade E, Blunsden S and Fisher R 2006 Modelling crowd scenes for event detection, *Pattern Recognition, ICPR 2006, 18th International Conference*.
- [22] Benezeth Y, Jodoin P, Saligrama V and Rosenberger C 2009 Abnormal events detection based on spatio-temporal co-occurrences, *CVPR'09: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, FL, USA, 2458–2465*.
- [23] Anti'c and Ommer B 2011 Video parsing for abnormality detection, *ICCV*.
- [24] Cheng K. W, Chen Y. T and Fang W. H 2015 Video anomaly detection and localization using hierarchical feature representation and gaussian process regression, *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [25] Li W, Mahadevan V and Vasconcelos N 2014 Anomaly detection and localization in crowded scenes, *Pattern Analysis and Machine Intelligence, IEEE Transactions*.
- [26] Laptev, Marszalek M, Schmid C and Rozenfeld B 2008 Learning realistic human actions from movies, *Computer Vision and Pattern Recognition, CVPR 2008, IEEE Conference*.
- [27] Shao, Loy C. C and Wang X 2014 Scene-independent group profiling in crowd, *2014 IEEE Conference on Computer Vision and Pattern Recognition*.
- [28] Saligrama V and Chen Z 2012 Video anomaly detection based on local statistical aggregates, *Computer Vision and Pattern Recognition (CVPR)*, 2012 IEEE Conference.
- [29] Xu, Ricci E, Yan Y, Song J and Sebe N 2015 Learning deep representations of appearance and motion for anomalous event detection, *BMVC*.
- [30] Yuan Y, Fang J and Wang Q 2015 Online anomaly detection in crowd scenes via structure analysis, *Cybernetics, IEEE Transactions*.
- [31] T. J. Vennila and V. Balamurugan, "A Rough Set Framework for Multihuman Tracking in Surveillance Video," in *IEEE Sensors Journal*, vol. 23, no. 8, pp. 8753-8760, April 2023.
- [32] Shehzed, Jalal A and Kim K 2019 Multi-Person Tracking in Smart Surveillance System for Crowd Counting and Normal/Abnormal Events Detection, *International Conference on Applied and Engineering Mathematics (ICAEM)*, 163-168.
- [33] Siqi Wang, En Zhu, Jianping Yin and Fatih Porikli 2017 Video anomaly detection and localization by local motion based joint video representation and OCELM, *Elsevier, Neurocomputing*, 1-15.
- [34] Vennila T. J and Balamurugan V "A Stochastic Framework for Keyframe Extraction," *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, Vellore, India, 1-5, 2020.
- [35] Wang, T. (2020). Spatio-Temporal Point Process for Multiple Object Tracking. *IEEE Transactions on Neural Networks and Learning Systems*. 34, 1777-1788.
- [36] <https://paperswithcode.com/dataset/crowd11>
- [37] <https://data.world/datasets/subway>
- [38] Ryan D, Denman S, Fookes C, Clinton B and Sridharan S 2011 Textures of optical flow for real-time anomaly detection in crowds, *2011 8th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 230–235.
- [39] Mehran R, Oyama A and Shah M 2009 Abnormal crowd behaviour detection using social force model, *IEEE Conference on Computer Vision and Pattern Recognition*, 935–942.
- [40] Mahadevan V, Li W, Bhalodia V and Vasconcelos N 2010 Anomaly detection in crowded scenes, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1975–1981*.
- [41] Wu S, Moore B.E, and Shah M 2010 Chaotic invariants of lagrangian particle trajectories for anomaly detection in crowded scenes, *Proceedings of the IEEE Computer Society Conference on Com Vision and Pattern Recognition, 2054–2060*.
- [42] Biswas S and Gupta V 2016 Abnormality detection in crowd videos by tracking sparse components, *Mach. Vision Appl.* 28, 35–48.
- [43] Lloyd K, Rosin P.L and Marshall D 2017 Detecting violent and abnormal crowd activity using temporal analysis of grey level co-occurrence matrix

(GLCM)-based texture measures, *Machine Vision and Applications* 28, 361–371.

- [44] Rajasekaran G and Raja Sekar J 2023 Abnormal Crowd Behavior Detection Using Optimized Pyramidal Lucas-Kanade Technique, *Intelligent Automation & Soft Computing*, 2439- 2412.
- [45] Md. Haidar Sharif, Lei Jiao and Christian W Omlin 2023 Deep Crowd Anomaly Detection by Fusing Reconstruction and Prediction Networks, *MDPI, electronics*, pp. 1-41.
- [46] Wang T and Snoussi H 2013 Histograms of optical flow orientation for abnormal events detection, *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS)*, 45–52.
- [47] Xu Y, Yan E, Ricci and Sebe N 2016 Detecting anomalous events in videos by learning deep representations of appearance and motion, *Computer Vision and Image Understanding*, 117–127.
- [48] Hasan, Choi J, Neumann J, Roychowdhury A. K and Davis L.S 2016 Learning temporal regularity in video sequences, *IEEE Conference on Computer Vision and Pattern Recognition*, 733–742.
- [49] Ionescu R. T, Smeureanu S, Alexe B and Popescu M 2017 Unmasking the abnormal events in video, *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy*, 2914–2922.
- [50] H. Song, C. Sun, X. Wu, M. Chen and Y. Jia, "Learning Normal Patterns via Adversarial Attention-Based Autoencoder for Abnormal Event Detection in Videos," in *IEEE Transactions on Multimedia*, vol. 22, no. 8, pp. 2138-2148, Aug. 2020.
- [51] G. Rajasekaran and J. Raja Sekar, "Abnormal Crowd Behavior Detection Using Optimized Pyramidal Lucas-Kanade Technique," *Intelligent Automation & Soft Computing*, pp. 2439- 2412, 2023.
- [52] Md. Haidar Sharif, Lei Jiao and Christian W. Omlin, "Deep Crowd Anomaly Detection by Fusing Reconstruction and Prediction Networks," *MDPI, electronics*, pp. 1-41, 2023.
- [53] S. Wang, Y. Zeng, Q. Liu, C. Zhu, E. Zhu, and J. Yin, "Detecting abnormality without knowing normality: A two-stage approach for unsupervised video abnormal event detection," in *2018 ACM Multimedia Conference, MM 2018, Seoul, Republic of Korea, October 22-26, 2018, 2018*, pp. 636–644



© 2024 by the Sujith Kumar P. S, Manish T I, Vince Paul and Sanaj M S. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).