IMPROVING LOCAL GOVERNMENT RESPONSE THROUGH AI-DRIVEN SPATIO-TEMPORAL CROWD DETECTION AND TRACKING

M.Mukesh Krishnan[1,*], S. Thanga Ramya[2], K. Ramar[3]

[1*]Assistant Professor, Department of Computer Science and Engineering. Francis Xavier Engineering College, Vannarpettai, Tirunelveli -627003, Tamilnadu, India.
[2] Professor, Department of Computer Science and Design, RMK Engineering College, Kavarapettai, Tiruvallur, Tamilnadu, India.
[3] Professor, Department of Computer Science and Engineering, RMK College of Engineering and Technology, Pudovoyal, Tiruvallur, Tamilnadu, India.

*mukeshkrishnan91.m@gmail.com[1]

Abstract

Unusual crowd behavior in public places can quickly turn into dangerous situations if it goes unnoticed. For local governments, being able to detect and respond to such events in time is vital for protecting citizens and maintaining public safety. Although many video-based systems have been developed over the years to monitor crowds, most are still limited in their ability to recognize abnormal activities in real-world, real-time conditions. To address this gap, our study presents a Spatio-Temporal Crowd Detection and Tracking (STCDT) approach that analyzes movement patterns to identify unusual behaviors. The method applies clear rules to detect events such as crowd merging, sudden running, or splitting, which often signal risk. By providing timely insights, this system can support local authorities in making faster and more effective decisions during emergencies. We test the approach on well-known datasets including UMN, Avenue, and UCSD, and the results show its strong potential to improve local government response in managing public spaces safely.

Keywords: Crowd Detection, Unusual Crowd Activities Detection, People Moving, and Spatio-Temporal Analysis

1. INTRODUCTION

Managing public safety in crowded spaces is a major concern for local governments, as unusual crowd behaviors such as sudden running, merging, or dispersal can quickly escalate into dangerous situations. Traditionally, surveillance cameras in public areas such as airports, educational institutions, train stations, markets, and stadiums have been monitored by human operators. However, continuous 24/7 monitoring is difficult, and lapses in concentration may lead to delayed recognition of abnormal events, creating serious safety risks [1]. This highlights the need for automated systems capable of detecting unusual events in real time.

In recent years, researchers have applied various artificial intelligence and pattern recognition techniques to video surveillance, aiming to identify unusual human actions and behaviors [2]. These systems define patterns through mathematical constraints and statistical measures to distinguish normal from abnormal activities. While progress has been made, traditional crowd detection methods struggle in dense environments where overlapping individuals and complex interactions reduce feature extraction efficiency [3–6]. This makes it challenging to identify unusual events accurately in real-world conditions.

To overcome these challenges, several models have been proposed, including particle filters, spatio-temporal gradients, and optical flow-based approaches [5]. Many studies have focused specifically on crowded scenes, analyzing motion patterns to differentiate between normal and abnormal activities [6–12]. However, crowd motion is often unstructured, which limits the reliability of these methods [13–16].

In this context, the present work introduces a Spatio-Temporal Crowd Detection and Tracking (STCDT) technique designed to enhance local government response in managing public spaces. The proposed approach continuously analyzes crowd dynamics using detection and tracking methods, estimates unusual crowd size by modeling relationships between adjacent groups, and applies rule-based mechanisms to identify key events such as merging, splitting, and sudden running. By providing real-time automated insights, this system reduces reliance on constant human monitoring and supports timely decision-making by local authorities, thereby improving public safety management.

The main contributions of this study are as follows:

1. Continuous analysis of spatio-temporal information for detecting unusual crowd activities.
2. Estimation of unusual crowd size through analysis of relationships between adjoining groups.
3. Definition of rule-based indicators for identifying unusual events such as crowd merging, splitting, and running.

2. RELATED WORKS

Most unusual crowd detection methods rely on pattern recognition techniques, where motion features are extracted to distinguish normal from abnormal behaviors. Several closely related works illustrate both the progress and limitations in this field. The closely related works are given here. Chunyu Chen and Yu Shao [17] developed a model that uses the energy of motion using optical flow. The energy of optical flow was defined in terms of weight and velocity. The weight is measured in how much variability between pixel coordinates in the consecutive frames. The velocity is measured by the ratio between horizontal direction coordinates and vertical direction coordinates of optical flow. In case of abnormal behaviours, the object movement and direction will be high. This model is constructed by consistent motion. In some situations, optical flow is unstable due to the quality of the video. It leads to poor accuracy because this model is worked by optical flow computation.

Siqi Wang et al. [18] created a normalcy model that uses Uniform Local Gradient Pattern-based Optical Flow (ULGP-OF) descriptor. In ULGP-OF, the gradient threshold was fixed by averaging eight surrounding pixels. If lesser than the threshold in the surrounding pixel it denotes '0' otherwise '1' and. The most two 0–1 or 1–0 jumps should be observed in one circular binary code. If more than two jumps are considered as drastic changes in the foreground, then, the histogram is formed and calculates the optical flow by estimating the bin. Though the man with a bike is normal speed, but that event is detected as abnormal event due to lack of structure of descriptor information.

Roberto et al. [19] developed an abnormality method by applying Foreground Occupancy (FO). FO is created by Gaussian Mixture Model (GMM) using cell-based, which is using mean, and standard deviation for finding active cells. For finding abnormal events, the posterior likelihood of the FO model is estimated. As a result, abnormal events are detected poorly due to sudden illumination changes in some environmental conditions. Although pattern matching based methods have proven successful in different applications, they still suffer from fundamental limitations [20-25] such as illumination variation, occluding the object. The pattern may not be matched since the abnormal is too small and covered by many surroundings pedestrians.In case of frames with low anomaly score, the pattern could not matched those frames correctly as abnormality. The abnormal event can be correctly detected only after the anomaly score increases and exceeds the threshold. As a result, the system cannot detect the actual frame of abnormal event start and end of abnormal event ends.

To address these shortcomings, researchers have explored learning-based methods that rely on inherent pixel-level characteristics [26–28]. These approaches extract features across entire frames, offering greater flexibility compared to handcrafted descriptors. However, even these methods must balance accuracy with

3. MATERIALS AND METHODS

3.1 Computing Infrastructure

The proposed STCDT method was implemented using Python 3.10 on a local machine equipped with Windows 11 Pro 64-bit operating system. The hardware configuration included an Intel® Core™ i7-12700K CPU running at 3.60 GHz, 32 GB of RAM, and an NVIDIA RTX 3080 GPU with 10 GB of VRAM. CUDA support was enabled to accelerate GPU-based operations and improve processing efficiency during video analysis tasks.

3.2 Dataset

1. UMN dataset: It is a benchmark dataset which contains 11 videos, the frame height is 480 and width is 640 (http://mha.cs.umn.edu/Movies/Crowd-Activity-All.avi). The videos are captured in sparsely crowded areas. All videos start with a normal crowd and after a certain period of time unusual crowd activities are there.

For example, people are walking gradually and later on this normal activity changes suddenly like people running [31].

2. Avenue dataset:This dataset consists of 37 videos which contains16 videos for training and 21 videos for testing [32]. There are 35240 frames for both training and testing videos where the frame height is 360 and width is 640 (https://www.cse.cuhk.edu.hk/leojia/projects/detectabnormal/dataset.html).

3. UCSD dataset: The dataset consists of two sub datasets namely Ped1 and Ped2. These datasets are captured in the people walking area (http://www.svcl.ucsd.edu/projects/anomaly/dataset.htm). This Ped1 dataset contains 70 videos and Ped2 dataset contains 28 videos. The ped2 dataset has 28 videos in which 16 videos for training process and 12 videos testing process. The following unusual activities such as people running, people merging, and vehicles appear in pedestrian paths in UCSD dataset [33].

3.2.1 Data Preprocessing

No extensive preprocessing was required for the datasets used in this study. All video frames were used in their original resolutions as provided by the respective sources. However, for computational efficiency and consistency, the videos were converted into frame sequences. Basic preprocessing steps included frame extraction, resizing (if required to match network input), and grayscale conversion to reduce computational complexity. No manual annotations, frame augmentations, or filtering were applied prior to processing with the STCDT pipeline.

3.3 Evaluation Method and Performance Metrics

To assess the effectiveness of the proposed STCDT method, a frame-level evaluation approach was adopted using three widely used benchmark datasets: UMN, Avenue, and UCSD (Ped1 and Ped2). The evaluation method involved testing the model within each dataset and comparing its performance with several existing approaches, such as GLCM, OPLKT-EMEHO, rpNet, and motion analysis techniques. This comparative analysis is presented in Tables 1 through 3, where the results of the proposed method are contrasted with those of existing techniques on the same datasets.

Along with the quantitative analysis, qualitative evaluation was also carried out by visually inspecting the output frames where unusual crowd events were detected. These results are illustrated in Fig. 2, which includes examples of crowd running, merging, splitting, and shaking, as identified by the proposed STCDT model

To evaluate the performance of the proposed STCDT, this paper used Receiver Operating Characteristic (ROC). To summarize the experimental results, Area Under Curve (AUC) is used. Further, the Equal Error Rate (EER) is used to estimate the equal probability for misclassifying frames where normal activity is classified as unusual crowd activity. The above metrics are used on the frame level criterion to evaluate unusual crowd activity scores of each frame.

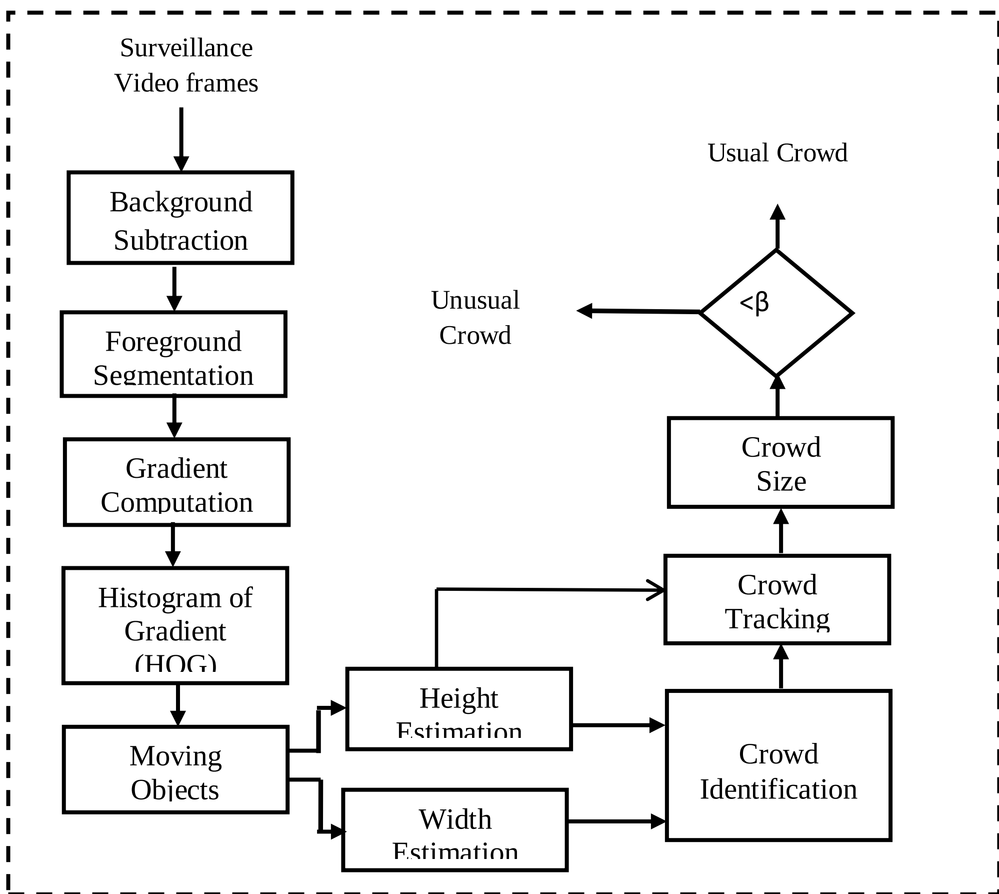
3.4 The Proposed STCDT

The aim of the proposed STCDT work is to detect the unusual moving crowd activities in the surveillance video by finding the relationship between detected objects. The STCDT consists of multiple stages namely background subtraction, foreground segmentation, gradient computation, Histogram of Gradient (HOG) estimation, moving objects identification, height and width estimation of moving objects, crowd identification and tracking, unusual crowd activities detection. The workflow of STCDT is explained in Fig. 1.

3.4.1. Problem Formulation

Let V is the surveillance video which consists of f_i frames where $i = 1$ to n that is $f_1, f_2, ...f_n$ is the individual frame. Let A_j be the number of foreground frames which is extracted from the background of the frame sequence where $j = 1$to n. Let B_k be the number of foreground objects in the A_j where $k = 1$ to m. Let P be the pixels that denote the number of pixels in both x and y directions. The P of B_k can express the size of the objects since the B_k may consist of one or more objects. To find the size of the foreground object in the B_k and to find the relationship among the detected and tracked objects, the spatio-temporal features are extracted on the B_j and matched it with the B_k in the consecutive frames. The size and distance of the B_k is determined whether it is an unusual crowd by applying STCDT.

Fig. 1 The workflow of STCDT



3.4.2. Foreground Objects Segmentation

The proposed STCDT considers N frames as an input. In general, the foreground objects contribute more details for unusual crowd detection and thus foreground objects need to be extracted in $f_1, f_2, ...f_n$. These frames are subtracted on the reference frame by using BM [29]to identify the foreground region. Let R be the reference frame, it is selected in the frame sequence, and it consists of only background objects. Each $f_1, f_2, ...f_n$ is subtracted from R. It is denoted using the eqn. 1.

$$P = \lfloor f_{xy} \rfloor - \lfloor r_{xy} \rfloor \qquad (1)$$

where f_{xy} and r_{xy} denote the corresponding position pixels of the x and y axis respectively.The resultant output frames are termed as A_j. To find the foreground objects easily, the A_j is converted into binary frames where foreground pixels are denoted as 1 value and background pixels are denoted as 0 value.

To improve the edge of the foreground object without any discontinuity, morphological operator is used where erosion technique is applied to get the connectivity if the foreground object breaks with multiple objects and dilation technique is applied to avoid unwanted regions of the object boundary. Next, connected components algorithm [30] is applied on the A_j to connect with upper, top, left, and right pixels foreground P by analysing foreground and background pixels. The output foreground objects are named as B_k in A_j.

3.4.3. Gradient Computation

To obtain the foreground objects spatial details, edge details are to be extracted in both x and y directions for controlling the degree of non-essential details of B_k. In this work, the sobel filter [18] is applied on P of the B_k by computing intensity values at each P. Firstly, the $A_j$is divided into non-overlapping multiple $\infty \times \infty$ regions with uniform size within the A_j. The edge details for both x and y directions are explained in the eqns. 2 and 3.

$$E_x = \{e_{x1}, e_{x2}, e_{x3}, ...e_{xn}\} \qquad (2)$$

$$E_y = \{e_{y1}, e_{y2}, e_{y3}, ..._{yn}\} \qquad (3)$$

whereE_x is the collection of x direction edges on B_k and E_y is collection of y direction edgeson B_k. The e_{x1}, e_{x2},...$e_{xn}$and e_{y1}, e_{y2},.....e_{yn} are the edges on x and y directions of each$P$.

3.4.4. HOG Estimation for Noise Removal

In noise removal stage, the E_x and E_y of B_k are used as the input for this stage. The edge extraction technique onthe foreground region in the previous stage might extract the HOG of the non-foregroundregion. Therefore, the extracted foreground regions have to be verified in the foreground detection stage. The foreground objects and non-foreground objects are varied by applying threshold value γ, it is explained in eqn. 4. If the foreground region is classified as a foreground object, it will be considered as the next stage, otherwise, it will be considered as a noise and it will be removed from the next process.

$$foreground \text{ object } = \begin{cases} \text{Yes,} & \text{if } B_k > \gamma \quad where\ k = 1\ to\ m \\ \text{No} & \text{Otherwise} \end{cases} \tag{4}$$

3.4.5. Moving Objects Identification

In the frame sequence, the output foreground objects B_k may or may not be moving. This work considers only moving objects because the moving object is highly involved in many unusual activities. The activities may be running, splitting, and merging from the crowd. In the work, the STCDT define the distance *(D)* between the corresponding moving objects B_k, it is computed between consecutive frames in terms of horizontal position x.Let $B_{1(A1)}$, $B_{2(A1)}, \ldots. B_{n(An)}$are the individual foreground objects of the A_j. The D is computed between corresponding detected $B_{1(A2)}$, $B_{2(A2)}, \ldots. B_{n(A2)}$ in the A_{j+1}. The D is computed using Euclidean distance measure; it is explained in the eqn. 5.

$$D\left(B_{k(A1)}, B_{k(A2)}\right) = \sqrt{\left(B_{kx(A1)} - B_{kx(A2)}\right)^2} \tag{5}$$

If D gets 0 value, the concerned foreground object is considered as static object in the frame, otherwise it is considered as moving objects.

3.4.6. Crowd Identification

To estimate spatial information of the B_k, the Height (H) and Width (W) are to be found using the foreground region of B_k. The position and size of the spatial information are needed to analyze how much area is covered by B_k. The H is computed in vertical direction *(y)* and W is computed in horizontal direction *(x)*. They are computed using the eqns. 6 and 7.

$$B_{k_H} = \sqrt{(B_{k\,y2} - B_{k\,y1})^2} \tag{6}$$

$$B_{k_W} = \sqrt{(B_{k\,x2} - B_{k\,x1})^2} \tag{7}$$

where B_{k_H} and B_{k_W} are the H and W of the foreground object respectively, $B_{k\,y2}$ and $B_{k\,y1}$ are the maximum and minimum vertical coordinates respectively, and $B_{k\,x2}$ and $B_{k\,x1}$ are the maximum and minimum horizontal coordinates respectively

In the unusual crowd activities analysis, the big size B_k is needed which should consist of more people that is called crowd and it is explained by using the eqn. 8. Let O_u be the number of crowds.

$$.crowd = \begin{cases} O_u, & if \;\; B_k > \alpha \\ B_k & Otherwise \end{cases} \qquad (8)$$

where α is the threshold value. It is applied on the B_k. If B_k size is greater than the α value, it is considered as the crowd, it is termed as O_u.

3.4.7. Crowd Tracking

The size of the O_u is continuously tracked in the consecutive frames to analyze whether it is involved in an unusual event or usual event. The O_u is tracked using its size using the eqns. 6. The size of the O_u is an important measure to track in the consecutive frames since the crowd size O_u is almost the same in the adjacent frame. In this work, the corresponding crowd is tracked in the consecutive frame. It is computed using the eqn. 9.

$$O_u = \begin{cases} yes, & \text{if} \;\; \left| \left(B_{k_H}\right)_{O_u}^{f_i} - \left(B_{k_H}\right)_{O_u}^{f_{i+1}} \right| < \chi \\ no, & \text{otherwise} \end{cases} \qquad (9)$$

In the corresponding O_u matching in the consecutive frame, if its height does not match, it is considered as O_u may change in the location order or O_u may merge with another O_u. To control the gradual movement variation, the threshold value χ is applied on the height variation between O_u in the consecutive frames.

G. Crowd Size Estimation for Unusual Crowd Activities Detection

The tracked O_u is continuously estimated its size to analyse whether it is unusual or usual. If the minimum and maximum are abruptly achieved from the size of O_u, it is called unusual crowd activities such as crowd running, crowd merging, and splitting. The following rules are defined to find out the unusual activities such as crowd merging, crowd splitting, and crowd running. The crowd merging activity is analyzed using eqn. 10.

$$unusual\ \text{crowd}\ merging = \begin{cases} \text{Yes}, & \text{if} \;\; O_u > \beta \\ \text{No} & \text{Otherwise} \end{cases} \qquad (10)$$

where β is the threshold value which is applied on the O_u. If O_u is greater than the β value, it is considered as the unusual crowd merging where the surrounding peoples merge together in the O_u. The crowd splitting activity is analyzed using the eqn. 11.

$$unusual\ \text{crowd splitting} = \begin{cases} \text{Yes,} & \text{if } O_u < \beta \\ \text{No} & \text{Otherwise} \end{cases} \quad (11)$$

If O_u is lesser than the β value, it is considered as the unusual crowd splitting where the crowd peoples split from the O_u. The crowd running activity is analyzed using the eqns. 12 and 13.

$$D\left(O_{u(A1)}, O_{u(A2)}\right) = \sqrt{\left(O_{ux(A1)} - O_{ux(A2)}\right)^2} \quad (12)$$

where $O_{u(A1)}$ is the crowd in the current frame and $O_{u(A2)}$ is the crowd in the next frame. The unusual moving is defined using the threshold value η.

$$unusual\ \text{crowd running} = \begin{cases} \text{Yes,} & \text{if } D\left(O_{u(A1)}, O_{u(A2)}\right) > \eta \\ \text{No} & \text{Otherwise} \end{cases} \quad (13)$$

3.5 Policy Relevance

While the STCDT method is technical in nature, its main contribution lies in enabling local governments and municipal authorities to respond to unusual crowd behaviors in real-time. By integrating such methods into public safety systems, local administrations can improve crowd management, disaster response, and urban governance efficiency.

4. RESULTS

In the UMN dataset, normal and unusual events frequently occur in indoor and outdoor scenes. The proposed STCDT detected unusual activity frames in appropriate places. Table 1 shows the unusual crowd detection results using UMN dataset and Table 2 show the performance analysis of STCDT on UMN dataset with several existing usual activities methods such as GLCM [34], OPLKT-EMEHO [35], rpNet [36], motion analysis [37].

TABLE 1 COMPARATIVE ANALYIS ON UMN DATASET

Scene No.	AUC (%)
Scene 1	98.47
Scene 2	97.17
Scene 3	98.42

TABLE 2 COMPARATIVE ANALYIS ON UMN DATASET

Methods	AUC (%)
GLCM	96.40
OPLKT-EMEHO	97.17
rpNet	96.21

Motion Analysis	96.47
STCDT	98.02

The quantitative nalysis is carried out using several existing unusual crowd activity methods and proposed STCDT using AUC and EER. Many unusual event detection research works have been taken the avenue dataset to detect the unusual crowd events because it includes several challenging scenarios. The proposed STCDT attained reasonable results on the Avenue dataset when compared with other existing methods. Table 3 shows the performance on Avenue dataset. Table 4 shows performance results on the avenue dataset.

TABLE 3 PERFORMANCE ANALYSES ON AVENUE DATASET

Video No.	AUC (%)	EER (%)
Video 1	97.12	17.18
Video 2	97.07	18.47
Video 3	96.47	17.02
Video 4	94.48	19.27
Video 5	92.17	19.21

TABLE 4 COMPARATIVE ANALYIS ON AVENUE DATASET

Methods	AUC (%)	EER (%)
Chong et. al [38]	81.07	19.08
Lonescu et. al [39]	79.09	-
Luo et. al [40]	81.00	-
Liu et. al [41]	82.15	-
Wang et. al [42]	87.14	20.12
Songet. al [43]	90.10	17.50
Rajasekaran and Raja Sekar [44]	91.30	-
Babu et al. [37]	93.44	-
STCDT	95.46	18.23

Table 5 and Table 6 show the performance on the UCSD dataset which consist of comparison results between proposed STCDT method and other existing methods in terms of AUC and EER. From this performance comparative analysis, the results are revealed that the proposed approach yields better results for both Ped1 and Ped2. The proposed STCDT correctly detected the unusual crowd activities such as crowd shaking and unusual crowds.

TABLE 5 PERFORMANCE ANALYIS ON UCSD DATASET

Sequence No.	UCSD (Ped1)		UCSD (Ped2)	
	AUC (%)	EER (%)	AUC (%)	EER (%)
Sequence 1	95.05	11.04	97.14	10.23
Sequence 2	96.03	12.01	97.00	12.42
Sequence 3	96.42	10.09	96.14	13.01
Sequence 4	94.48	11.42	95.40	11.53
Sequence 5	95.44	12.06	94.33	11.18

TABLE 6 COMPARATIVE ANALYIS ON UCSD DATASET

Methods	UCSD (Ped1)		UCSD (Ped2)	
	AUC (%)	EER (%)	AUC (%)	EER (%)
Xu et. al [50]	90.08	14.70	90.00	15.20
Hasan et. al [51]	87.21	24.00	92.11	21. 70
Lonescu et.al [39]	73.00	-	80.10	-
Chong et. al [38]	89.40	10.80	85.67	10.70
Liu et. al [41]	84.00	-	93.10	-
Wang et. al [42]	80.15	33.27	95.44	08.16
Song et. al [43]	90.10	15.00	90.00	15.15
Rajasekaran and Raja Sekar [44]	90.10	14.00	91.10	16.70
Haidar et al. [45]	90.50	13.00	92.20	10.10
Babu et al. [37]	92.14	12.47	94.00	10.17
STCDT	95.48	11.32	96.00	11.67

Fig. 2 The output unusual event frames on UMN dataset, Avenue, and UCSD datasets

The STCDT method demonstrates the qualitative results and it shows in Fig. 2 where the figure consists of all detected experimental positive results in all three datasets. The detected unusual crowd activities are crowd running and unusual crowd in UMN dataset, unusual crowd running, crowd merging in Avenue dataset, and crowd splitting and crowd merging in the UCSD dataset.

In addition, the STCDT revealed that the higher AUC values and EER values in all three datasets where 98.02% (AUC) for UMN dataset, 95.46% (AUC) and 18.23% (EER) for avenue dataset, and 95.48% (AUC) and 11.32% (EER) for UCSD Ped1 dataset and 96.00% (AUC) and 11.67% (EER) for UCSD Ped2 dataset. The proposed STCDT detected more unusual crowd events in the UMN dataset than the UCSD dataset. Therefore, the UMN dataset obtained higher AUC than the other two datasets and the UMN results are shown in Table 2. In Avenue and UCSD dataset, many unusual crowds and unusual crowd running are detected by the proposed method and there are no other unusual crowd activities detected such as crowd merging and splitting. The proposed STCDT works well in case activities change abruptly in the consecutive frames since the STCDT computes the distance between the corresponding crowds in the consecutive frames. In some cases, the unusual event occurs gradually in the consecutive frames. In this case, the algorithm gives low detection accuracy.

The STCDT method consistently outperformed other state-of-the-art techniques across all three benchmark datasets. It achieved an AUC of 98.02% on the UMN dataset, 95.46% on the Avenue dataset, and 95.48% and 96.00% on UCSD Ped1 and Ped2, respectively. These results highlight the method's reliability in detecting both sudden events, such as crowd running, and more complex scenarios, including crowd merging and splitting. Beyond its technical performance, this level of accuracy has important governance implications. Reliable detection of unusual crowd behavior can help local governments and public safety authorities respond more quickly to potential threats in public spaces. When integrated into existing surveillance systems, STCDT offers a practical tool to support better crowd management, strengthen emergency preparedness, and enhance overall public safety.

5. CONCLUSION

This paper introduced STCDT method using moving human detection and tracking techniques by applying spatio-temporal features to detect unusual crowd event detection. To effectively detect the unusual crowd events, the STCDT is built using multiple stages such as background subtraction, foreground segmentation, gradient computation, Histogram of Gradient estimation, moving objects identification, crowd identification and tracking, and unusual crowd activities detection. These combinations of STCDT detected many unusual crowd events such as unusual crowds, unusual crowd merging, and unusual crowd splitting by applying multiple unusual crowd events rules. The experimental results are carried out on three challenging datasets and the results indicate the effectiveness of the proposed STCDT method by comparing several existing methods. While the STCDT method performs well in detecting a range of unusual crowd activities, a few limitations to note. Variations in camera angles, lighting conditions, and occlusions can affect detection accuracy. This limitation provides directions for future research, such as integrating learning-based techniques and domain adaptation strategies to enhance robustness and scalability. From a governance perspective, the ability to detect unusual crowd behaviors in real time can significantly support local governments and municipal authorities in managing public safety, improving emergency preparedness, and ensuring more effective crowd control in urban spaces.

REFERENCES

1. Asha Paul, M. K., Kavitha, J., & Jansi Rani, P. A. (2018). Key-Frame Extraction Techniques: A Review. Recent Patents on Computer Science, 11(1), 3–16.
2. Ang Li, Zhenjiang Miao, Yigang Cen, Xiao-Ping Zhang, Linna Zhang, and Shiming Chen, "Abnormal event detection in surveillance videos based on

low-rank and compact coefficient dictionary learning," Elsevier, pattern recognition, pp. 1 16, 2020.

3. E. Murali, A.C. Santha Sheela, M. Asha Paul, V. Muthu, A. Yovan Felix, "Learning normal patterns via conv-LSTM for video anomaly detection using likelihood statistical texture feature representation in surveillance videos." International Journal of System Assurance Engineering and Management, Vol. 16, 2025, pp. 1-12.
4. Kavitha, J., P. Arockia J. Rani, P. Mohamed Fathimal, and Asha Paul. "An Efficient Shot Boundary Detection Using Data-cube Searching Technique." Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science) 13, no. 4 (2020): 798-807.
5. N.N.A. Sjarif , S.M. Shamsuddin , S.Z. Hashim , Detection of abnormal behaviors in crowd scene: a review, Int. J. Advance. Soft Comput. Appl. 4 (1) (2012) 1–33 .
6. Peer Mohamed Appa, M. A. Y., Vanitha, V., Rishi, P., Sagar, S., & Paul, M. A. (2025). Key frame extraction based abnormal vehicle identification technique using statistical distribution analysis. Scientific Reports, 15(1), 30957.
7. C. Ma , Z. Miao , X. Zhang , M. Li , A saliency prior context model for real-time object tracking, IEEE Trans. Multimed. 19 (11) (2017) 2415–2424 .
8. J. Geng , Z. Miao , X. Zhang , Efficient heuristic methods for multimodal fusion and concept fusion in video concept detection, IEEE Trans. Multimed. 17 (4) (2015) 498–511
9. Asha Paul, M., Jansi Rani, P. A., & Liba Manopriya, L. (2020). Gradient based aura feature extraction for coral reef classification. Wireless Personal Communications, 114(1), 149–166.
10. M. Rodriguez, S. Ali, T. Kanade, Tracking in unstructured crowded scenes, in: Proc. IEEE Int. Conf. Comput. Vis., 2009, pp. 1389–1396.
11. Y. Cen, Y. Cen, K. Wang, et al., Energy-efficient nonuniform content edge pre– caching to improve quality of service in fog radio access networks[J], Sensors 19 (6) (2019)
12. H. Wang, Y. Cen, Z. He, et al., Robust generalized low-rank decomposition of multimatrices for image recovery[J], IEEE Trans. Multimed. 19 (5) (2016) 969–983.
13. H. Wang, Y. Cen, Z. He, et al., Reweighted low-rank matrix analysis with structural smoothness for image denoising[J], IEEE Trans. Image Process. 27 (4) (2017) 1777–1792.
14. H. Wang, Y. Li, Y. Cen, et al., Multi-Matrices low-rank decomposition with structural smoothness for image denoising[J], IEEE Trans. Circuits Syst. Video Technol. (2019).

15. Vennila TJ and Balamurugan V, “A stochastic framework for keyframe extraction,” 2020 IntConfEmerg Trends Inform TechnolEng (ic-ETITE), pp. 1–5. 2020.
16. Li, Z. Miao, Y. Cen, Global abnormal event detection based on compact coefficient low-rank dictionary learning, in: Asian Conf. Pattern Recognit. (ACPR), 2017, pp. 483–487.
17. Duan-Yu Chen & Po-Chung Huang, 2011, ‘Motion - based unusual event detection in human crowds’, Elsevier, J. Vis. Commun. Image Representation, vol. 22, pp. 178 – 186.
18. Siqi Wang, En Zhu, Jianping Yin &FatihPorikli, 2016, ‘Video anomaly detection and localization by local motion based joint video representation and OCELM’, Elsevier, Neurocomputing, vol. 277, pp. 161 – 175.
19. Roberto Leyva, Victor Sanchez & Chang - Tsun Li, 2017, ‘Video anomaly detection with compact feature sets for online performance’, IEEE Transaction on Image Processing, vol. 26, pp. 3463 – 3478.
20. Shehzed, A, Jalal, A & Kim, K 2019, ‘Multi-person tracking in smart surveillance system for crowd counting and normal/abnormal events detection’, in Proc. Int. Conf. Appl. Eng. Math. (ICAEM), pp. 163–168.
21. Takahashi, M, Fujii, M, Shibata, M & Satoh, S 2010, ‘Robust recognition of specific human behaviors in crowded surveillance video sequences’, EURASIP J. Adv. Signal Process., vol. 2010, no. 1, pp. 1–14.
22. AshishSinghal, Pradeep Kumar, RajkumarSaini, ParthaPratim Roy, Debi ProsadDogra&Byung-Gyu Kim, 2018, ‘Summarization of videos by analyzing affective state of the user through crowdsource’, Elsevier, Cognitive Systems Research, vol. 52, pp. 917–930.
23. Chen, C & Shao, Y, 2014, ‘Anomalous crowd behavior detection and localization in video surveillance’, 2014 IEEE International Conference on Control Science and Systems Engineering, Yantai, pp. 190-194.
24. GuogangXiong, Jun Cheng, Xinyu Wu &Yen-Lun Chen, YongshengOu&YangshengXu, 2012, ‘An energy model approach to people counting for abnormal crowd behaviour detection’, Elsevier, Neurocomputing, vol. 83, pp. 121-135.
25. Shehzed, A, Jalal, A & Kim, K 2019, ‘Multi-person tracking in smart surveillance system for crowd counting and normal/abnormal events detection’, in Proc. Int. Conf. Appl. Eng. Math. (ICAEM), pp. 163–168.
26. Takahashi, M, Fujii, M, Shibata, M & Satoh, S 2010, ‘Robust recognition of specific human behaviors in crowded surveillance video sequences’, EURASIP J. Adv. Signal Process., vol. 2010, no. 1, pp. 1–14.
27. AshishSinghal, Pradeep Kumar, RajkumarSaini, ParthaPratim Roy, Debi ProsadDogra&Byung-Gyu Kim, 2018, ‘Summarization of videos by

analyzing affective state of the user through crowdsource', Elsevier, Cognitive Systems Research, vol. 52, pp. 917–930.

28. Chen, C & Shao, Y, 2014, 'Anomalous crowd behavior detection and localization in video surveillance', 2014 IEEE International Conference on Control Science and Systems Engineering, Yantai, pp. 190-194.
29. Elhoseny M, "Multi-object detection and tracking (MODT) machine learning model for real-time video surveillance systems," Circuits systems and signal processing. Springer, pp 611–630, 2020.
30. Vennila TJ, Balamurugan V, "A rough set framework for multihuman tracking in surveillance video," IEEE Sensors, 2023.
31. umn-dataset-anomaly-in-crowd_fg3_270790763
32. CUHK Avenue Dataset | Papers With Code
33. UCSD Anomaly Detection Dataset
34. 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.
35. RajasekaranG and Raja Sekar J 2023 *Abnormal Crowd Behavior Detection Using Optimized Pyramidal Lucas-Kanade Technique,*Intelligent Automation & Soft Computing, 2439- 2412.
36. 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.
37. K. Babu, L. Godlin Atlas, S. Sreeji, K. P. Arjun,· A. Rasmi, and Gururama Senthilvel, "Appearance and motion-based unusual crowd events detection using multiple moving objects," multimedia and applications, pp. 1 – 13, 2024.
38. Chong Y. S and Tay Y. H 2017 *Abnormal event detection in videos using spatiotemporal autoencoder,* 189–196.
39. 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.
40. Luo W, Wen L and Gao S 2017 *A revisit of sparse coding based anomaly detection in stacked rnn framework*, IEEE International Conference on Computer Vision.
41. Liu W, Luo W, Lian D and Gao S 2018 *Future frame prediction for anomaly detection - A new baseline*, 2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City.
42. Wang S, Zeng Y, Liu Q, Zhu C, Zhu E and Yin J 2018 *Detecting abnormality without knowing normality: A two-stage approach for unsupervised video abnormal event detection,* 2018 ACM Multimedia

Conference on Multimedia Conference, MM 2018, Seoul, Republic of Korea, 636–644.

43. Song H, Sun C, Wu X, Chen M and Jia Y 2020 *Learning Normal Patterns via Adversarial Attention-Based Autoencoder for Abnormal Event Detection in Videos, IEEE Transactions on Multimedia*, 2138-2148.
44. RajasekaranG 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.