July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

Integrated Detection for Behavior Recognition in Videos

Savitha Chikkathimmaiah[1], Dodderi Ramesh[2]
[1]Department of ECE, Sri Siddhartha Institute of Technology Tumakuru, Sri Siddhartha Academy of Higher Education, Karnataka, India, savithac1122@gmail.com
[2]Department of CSE, Sri Siddharth Institute of Technology Tumakuru, Sri Siddhartha Academy of Higher Education, Karnataka, India, rameshd_ssit@gmail.com

Article Info
Volume 83
Page Number: 6320 - 6336
Publication Issue:
July - August 2020

Article History
***Article Received**: 25 April 2020*
***Revised**: 29 May 2020*
***Accepted**: 20 June 2020*
***Publication**: 10 August 2020*

Abstract

In computer vision, activity recognition in video has attracted researchers due to its variety of applications such as, human computer interaction, video retrieval and surveillance system. At the same time it's really challenging to detect actions in real-time world, due to its complex motion style and background litters. This causes several confusion. Videos of high-dimensionality also limit the performance of recognition. Many number of features are required in order to obtain good action representation and also to minimize asynchrony amongst stream data. The feature of Motion Heat Map (MHM)is considered to represent group activity and implement motion information to signify the trajectory. Global motion pattern information is obtained using optical flow. The gained feature vectors from optical flow and MHM are combined and fed as input to bag of words approach in order to detect normal and abnormal frames. For simulation and validation, widely used dataset like UMN is considered to validate the proposed model.

Keywords: Bag of Words (BoW), Motion Heat Map (MHM), Optical Flow (OF), Features Integration, Recognition Rate (R-Rate).

1. INTRODUCTION

The anomaly of crowd behavior is an interesting event in a crowded environment such as public places. The smart vision should transmit and encode the scene of a video if any abnormal event is identified in the same scene. Any anomalous behavior or speed in individuals or crowd motion is considered an event that can be represented using trajectory. Moreover in scene, new object appearance constitutes an event. The sensor capable of recognizing the anomalies in crowd-behavior is transmitting and encoding the video data. It can significantly minimize computational overhead and thus the necessity of the manpower. Additionally, the identification of automatic events in the crowd can lead to quick response from concerned personnel. In the last few decades, the research in computer vision has been accelerated. Several computer vision approaches have been introduced for a

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

variety of applications. These algorithms can be performed in the vision sensor to recognize interest events. The identification of an interesting event can improve the efficiency of the video processing system. In other words, the computational resources are saved by the help of processing data, which consists of an interesting event [1]. The identification of crowd behavior is most challenging but also very eminent for computer-vision applications. The automatic identification of anomalies in crowd behavior will not only minimize the needed number of human monitors but will facilitate a very quick response to calamities.

One of the most common solutions to these issues is to discover novel patterns in data that do not ensure expected case [2]. This can be improved by fitting the statistical model of anomaly identification that tries to recognize events with lower probability as an abnormality. Anyways, it represents many number of challenges. In first, it requires the higher dimensional feature to represent better event and train the statistical model. In this case, several training samples will improve exponentially with the help of the feature dimension. In practice, it is very difficult to gather enough data to train the statistical model. Secondly, the crowded scene needs statistical identification model that is robust to the scenes that are dynamic and complex and consists of a large number of moving people who conclude each other in a complex way and have a lower resolution. Henceforth, it is very difficult to recognize all abnormal behaviors. Lastly, various tasks may need various normalcy models. So, it is not realistic to gather sufficient samples of abnormal video outcomes that bring the challenges to create a more robust video anomaly detector [3] [4].

The utilization of mid-level and low-level crowd features may generate more difficulty at the identification of a higher-level of crowd behavior from publicly holistic sight. In order to get more insight into information, many methods can be utilized as group structure [5], energy potentials [6], streak line potentials [7], Spatio-temporal and motion heat-map. Streak-line is utilized to recognize the temporal and spatial differences inflow, which can be obtained via the time-integration of the velocity-field [7]. The potential functions can be gained from spatial integration and each of the integration gives valuable data regarding changes in the scene. Research analysis of social behavior in the crowd scenarios reveals a public trend to follow the trail walkers who have the same group path. As the walker passes from the point, there has been social expectation that another walker behind them to follow the same path. In order to accumulate all social behaviors, the gap between actual flow and optic flow should be occupied with the same motion vector through all trajectories to analyze the crowd-behavior.

Here the main focus is on key problems that generally happen in the surveillance system and also their effects on performance. Here, many number of features are used in order to obtain good action representation and also to minimize asynchrony amongst stream data to obtain the optimal motion-wise complementary data. So good features are required to identify the action. The features like MHM are considered to

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

represent group activity and abnormal individuals.MHM can efficiently recognize many activities in motion. Then implement this motion information to signify the trajectory and global motion pattern. Information obtained is appropriately modeled using optical flow. The gained feature vectors from optical flow and MHM are combined to find-out abnormal frame-scene. UMN dataset is widely considered for the research purpose and to simulate the analysis of the result.

2. LITERATURE SURVEY

Several researchers have worked on the identification of crowd anomaly in real-life demands and have represented [8] a detailed survey of the identification of the anomalous events. In past decades, various methods have been used to identify the identification of anomalous crowd. Few of them are described as follows. Crowd is a collection of individuals. Few researchers have modeled the individual crowd in the scene. If the individual behavior is different from learned behavior then the anomaly is recognized. Whereas in [9], it utilizes unsupervised learning to learn individual behavior in the scene. Afterwards, utilized the likelihood ratio of individual behavior against the normal behavior to define the anomaly.

Similarly in paper [10], the individual-behavior of the atomic events is modeled in the scene. In order to define anomaly, the mixture of atomic activities have been compared against three normal classes. Recognizing individual-behavior is not the common method to recognize anomalous crowd behaviors. Several researchers have assumed the crowds holistically and then model the crowd-behavior rather than the individuals. In this method [11], the author modeled the crowd-behavior based on actual and desired velocities which are obtained via the particle advection. Whereas in paper [12], the crowd-behavior is modeled utilizing the potential of the interaction potential. They utilize the interest points of space-time [13] and trail them to the model of human motion in the video.

The interaction-energy represents whether the objects will meet in upcoming videos or not, by utilizing the interest points of space-time and then utilize the energy to recognize the anomalous events. To recognize the abnormal crowd-behavior spatially, local identification of the anomalous activity has been utilized. Such methods utilize objects behavior in pixel's neighborhood by the local-feature. The salient features are utilized to define anomalous events. In paper [14], it utilizes the bottom-up saliency to recognize the global-rarity of local-features that is composed of various speeds and directions of object move. A heat map motion is enhanced to compare the local-motion against all motions in this scene. Whereas in [15], it represents a novel feature map named as MDT (Mixture-Dynamic-Texture). The MDT assumes both motion and appearance-based object-properties in the scene. The MDT is utilized to model the object's moving behavior in pixel's' neighborhood. Temporal as well as spatial saliency is assumed when defining the anomalous event. The expansion over MDT has been introduced in [16] that

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

utilizes the CRF (Conditional-Random-Field).

In paper [17], the descriptor-based HOFM (Histogram of Optical-Flow-Orientation) is utilized and categorize the anomalous by utilizing single-class Support vector Machine. Few researchers have modeled the crowd behavior in fixed-cuboids. The cuboids have fixed-position and the windows are cropped from a consecutive set of a video frames that are stacked together. Whereas in [18], the crowded sequences are analyzed and encode the cuboids motion-patterns using HMM (Hidden Markov Models). In [19], the intensity can be measured over the cuboid utilizing wavelet transform that changes with frequency. The abnormal event represents high frequency under the time claimed by the author. In [20], the author performed a Spatio-temporal analysis of video segments in order to extract the descriptors based on the multi-scale OFH (optical-flow-histogram). The descriptors are utilized for recognizing the anomalous event.

In paper [21], the track lets statistics are estimated to model normal crowd-behavior in fixed cuboids. The advanced version of iHOT (improved Histogram of Track lets) has been represented in [22]. The sparse dictionary-based approach has tempted the researcher interest [23], [24]. In [23], the author does not encode the entities of the dictionary but rather utilizes the gradients of Spatio-temporal that directly populate the dictionary. Resulting at the frame rate of 150fps. Equally, the codebook model is utilized to recognize the anomalies in [25]. Whereas in [26], the cost of sparse reconstruction is described over normal events-dictionary to define the test sequence. Similarly in [27], the 2 parts of the dictionary are utilized for recognizing the dictionary features in cuboids to recognize the anomalous behavior. Presently, DL (Deep Learning) methods have been represented. In [28], the author considered Spatio-temporal context as dictionary-features in cuboids to recognize anomalous behavior. Presently, DL based methods have been introduced. In [29], utilizes ML (Machine Learning) framework to learn the features from video rather than utilizing the handcrafted features. The models of bag of words is utilized to score the anomaly level of each input from learned features. The outcomes of bag of words are used for the final anomaly identification.

3. PROPOSED METHODOLOGY

Finding behavior changes is a very difficult process in dense crowd scenario, where the context scene and dynamics of crowd change in an extensive range. In this paper, fusion approach is proposed by taking feature vectors from motion heat map (MHM) and optical flow (OF). The obtained features represents the individual/group activity. The MHM is efficient in tracing the motion information. The process of thermal energy distribution is considered to develop heat-map (HM). This is very useful to capture individual’s motion uncertainty. Figure 1 shows the block diagram of proposed fusion approach.

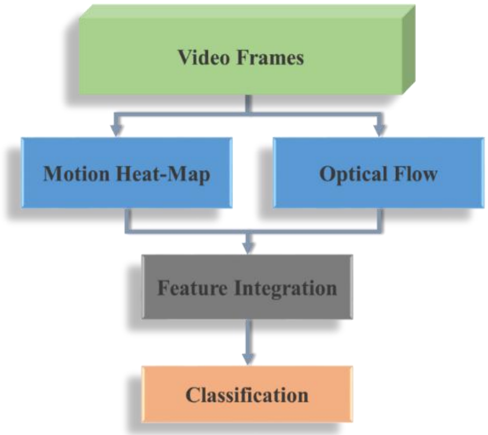

Figure 1:Flow diagram of Proposed Model

3.1 Feature Extraction using MHM

The motion information acquired directly from the global feature vectors may lose some of the important information, therefore in order to avoid this consider the heat source series as active trajectories. In addition, decay factor has been used on some heat sources so the previous heat sources will become lesser important as per the greater thermal energies of new heat sources. However, the heat source thermal values can be arranged in accordance to trajectories direction and temporal information. Several studies at trajectories has indicated larger differences, therefore usage of heat source as a feature can affect the motion fluctuation. In order to minimize the fluctuation during motion, consider the distributed process of thermal energy to diffuse originated heats by heat source series at all scenes. The outcome of diffusion tends to motion called MHM and features of MHM provides the motion activity information through 3D surfaces.

Here, x denotes the total number of trajectories in a group activity and A_y signifies the thermal energy at considered patch y. Then heat source thermal energy can be computed as,

$$A_y = \sum_x \bar{A}_{y,x} \cdot e^{-B_a\left(a_{aj} - a_{qs,x}\right)} \qquad (1)$$

The exponent value $e^{-B_a\left(a_{aj} - a_{qs,x}\right)}$ signifies time decay [30], whereas B_a denotes the coefficient value for temporal decay. $a_{qs,x}$ is frame number when x^{th} trajectories leaves patch y. $\bar{A}_{y,x}$ gives thermal energy accumulated for trajectory x in patch y and is given by,

$$\bar{A}_{y,x} = \int_0^{a_{qs,x} - a_{ds,x}} D\, e^{-B_a a}\, dt \qquad (2)$$

After performing the integration, equation (2) can be written as,

$$\bar{A}_{y,x} = \frac{D}{B_a}\left(1 - e^{-B_a\left(a_{qs,x} - a_{ds,x}\right)}\right) \qquad (3)$$

Where, $a_{ds,x}$ and $a_{qs,x}$ representsthe considered frame at epoch, when x^{th} trajectory seems with the patch y, B_a in equation(1) represents the coefficient of temporal decay. However, the novel heat trajectory sources require supplementary thermal energies when compared to the preceding heat sources. D is a constant set as 1 to obtain the thermal energy of equation (3) which is proportionate to trajectories length x at patch y.

At some particular instance, if x is continued in y for a long period, it will continue to collect added thermal energy at y patch. In different scenarios, when the availability of trajectory is not there at patch y, zero value is considered to be accumulated thermal energy which signifies y patch unavailability as the HM patch. Afterwards, the distribution process of thermal energy is performed to

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

obtain HM of scene and HM value at patch y is computed by,

$$E_y = \frac{\sum_{i=1}^{F} G_i \cdot e^{-A_{sdc}\, diff\,(y,i)}}{F} \quad (4)$$

For patch i, the thermal energy source is specified by G_i and F. This denotes a complete patch of heat source. The spatial diffusion coefficient is represented as A_{sdc}. The distances between y and i patches are given as $diff(y,i)$.

The acquired highest peak point h and the points of peak location is used as input to MHM and alignment steps are represented as $(K_1, K_2, K_3, K_4 \ldots, K_h)$. The pointed position K_y is provided by $K_y = [m_y, n_x]$ at coordinates m and n in MHM, where the K_y organization generally depends on calculated heat values. Scaling point at the MHM is given as,

$$\frac{\sum_{y=1}^{h} \sqrt{m_y^2 + n_y^2}}{h} = 1 \quad (5)$$

The interest point alignment with thebesiegedHM can be given as,

$$\arg_L \left(\min \sum_{y=1}^{h} \left| P_y - K_y . L \right|^2 \right) \quad (6)$$

Where, the interest point in K_y is L. At express interest points for the HM target.

However, alignment matrix is also noted for HM that it is lacking for trajectories. So the patches with respect to measured motion patches will become source to produce HM. Therefore the trajectories reliability becomes very difficult to achieve. So it is essential to put on motion process of low-level to produce HMs features as recognition process followed as optical flow [31].

3.2 Global Motion Pattern recognition

Entire motion information is necessary for the trajectory, and global motion pattern is required for a scene to model appropriately. Q trajectories have been considered from a standing scene and is denoted by $\{\alpha^q\}_q^Q = 1$, also an individual trajectory can be calculated as,

$$\alpha^q = \{\alpha_h^q\}_{H=1}^{\mathrm{Tr}^q} \quad (7)$$

Here, $\alpha_h^q \in T$ signifies the h^{th} position trajectory point, α^q trajectory length is denoted through Tr^q and the motion at particular position point is computed as,

$$r_h^q = \frac{K_{h+1}^q - K_h^q}{\Delta \mathrm{T}^{sp}} \in T \quad (8)$$

Sampling period among the adjacent points are provided as T^{sp} and measured trajectories Q is modeled in order to become optical flow in a particular scene. An individual K_H^Q represents a sample flow and generally corresponds towards r_H^q. Thisrefers to optical flow movement and results in energy HP transfer at point K_H^q.

Suppose trajectories link have an unchanged flow, then at this period the optimal thermal energy is accumulated in field S byminimizing the entire thermal energy. Afterwards it is being transferred at particular time interval in the process of optical flow.

$$\frac{min}{S} = \sum_{rj \in rj} \sum_{K \in f} \Delta G\big(s(K, rj)\big) \quad (9)$$

TEST Engineering & Management

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

Where,

$$s(K,rj) \geq 0 \ \ and \ \sum_{K \in f} s(K,rj) = s, \tag{10}$$

The acquired thermal energy from position K atany time period is considered as $\Delta G\big(s(K,rj)\big)$. Where,$(s > 0)$is defined to be constant. Eq. (9) and (10) have constraint value and acquire the suitable circulation of $s(K,rj)$so as to avoid infinity values. The values of $\Delta G\big(s(K,rj)\big)$can be computed by the process of thermal transmission [32].

$$\Delta G\big(s(K,rj)\big) = M \frac{u(K)r(K,rj)}{s(K,rj)} \tag{11}$$

M shows a parameter that is associated with heat alteration between the Kth position and their neighbors [32]. $u(K)$express density flow at position K. $r(K,rj)$signifies moving velocity corresponding to Kth position direction.

Commonly, $\Delta G\big(s(K,rj)\big)$definestotal thermal energy received through their K neighbor locations.ΔG is amount of energy completely transferred from the Kth position.

During generation of equipotential lines, primarily create thermal diffusion map on K_H^qinput trajectory points by dispensing thermal energies to scene S.Afterwards, energy constant line through energy map is used to apprehend all related information of motion at the considered trajectory points. Each trajectory point atK_H^qmeans its consistent of thermal diffusion map.

$$G\big(K_H^q\big) = \big[G\big(K_H^q, V\big)\big]_{Wid \times Hit}, K \in f \tag{12}$$

The thermal energy at Kth positioninK_H^qmaps is represented by $G\big(K_H^q, V\big)$, where Hit and Widexpress the scene 'height and width'. Scene grid positioning is given by,

$$f = [1, \ldots, Hit] \times [1, \ldots, Wid] \tag{13}$$

In time function consider $G\big(K_H^q\big)$. Therefore computation of $G\big(K_H^q\big)$by an iterative process is possible and initial map$G_0\big(K_H^q\big)$is given by,

$$G_0\big(K_H^q, K\big) = \begin{cases} G \text{ if } K = K_H^q \\ 0 \text{ otherwise} \end{cases} \tag{14}$$

The primary 'thermal energy' is represented by$G_0\big(K_H^q, K\big)$at position K and their 2nd order derivative function is written as,

$$\frac{\partial G\big(K_H^q, K\big)}{\partial a} = \sum_{rj \in Rj} s(K,rj) \frac{\partial^2 G\big(K_H^q, K\big)}{\partial rj^2} \tag{15}$$

While considering above equation (15), the process of thermal diffusion at their neighbor position for a solitary iteration is computed as,

$$W_T\big(V_G^k, V\big) \approx \frac{\sum_{K' \in N(K)} G_{Z-1}\big(K_H^q, K'\big) \times \left(-\frac{\|K-K'\|}{s \times \left(K, \frac{(K-K')}{\|K-K'\|}\right)}\right)}{|N(K)|} \tag{16}$$

Here, $N(K)$is for K neighborhood, and the neighborhood size is denoted by$|N(K)|$. While,$K' \in N(K)$represents neighbor of K, distance of neighbors is calculated as$\|K - K'$.The thermal energy coefficient at K position is specified by $s \times \left(K, \frac{(K-K')}{\|K-K'\|}\right)$.

The efficient optical flow is provided by considering center point A_k and the points of boundary are set as $(A_y)_y = [1, \ldots, W_A]$. I_D signifies total measured boundary points. The computed distance A_k and A_y represents relative boundary of the decision point. Distance between A_k and A_y on a time-axis is given by C_y^a. To simplify computation process, C_y^a is calculated in an output spatial plane.

This process allows A_k to follow the input route of trajectory and is indicated as $\{K_{k,h}\}_{H=1}^{\mathrm{Tr}} = 1$, where Tr denotes input trajectory length and H^{th} input trajectory point location at the spatial temporal provided by $K_{k,h}$. Congruently, the boundary point is represented by A_y (such as, $y \in [1, \ldots, W]$). In addition, the constructed path of y^{th} boundary pointsis given by $\{K_{y,H}\}_{H=1}^{\mathrm{Tr}}$. Thelocatedposition on $b^{\mathrm{th}} K_{k,h}$ is $K_{y,H}$. Likewise, on time-axis the concealed distance is given by C_y^a of A_k and A_y.

$$C_y^a \propto \frac{1}{\mathrm{Tr}} \sum_{H=1}^{\mathrm{Tr}} (\beta_{k,H} - \beta_{y,H}) \tag{17}$$

Where $\beta_{k,H}$ and $\beta_{y,H}$ denotes time-axis velocity, when the value of A_k and A_y pass through location $K_{k,H}$ and $K_{y,H}$. Moreover, to normalize C_y^a, consider $1/\mathrm{Tr}$. $\beta_{k,H} - \beta_{y,H}$ in eq. (17) isvalued as,

$$\beta_{k,H} - \beta_{y,H} \approx \left(1 - \frac{\gamma_1}{\mu_{y,H-1}^T}\right)(Rk - R_{y,H} - 1) + \gamma_2 cos\theta_{y,H-1} \tag{18}$$

The values of γ_1 and γ_2 are considered to be constant coefficients. Velocity at initial point is considered to be zero in the process. Here, $\mu_{y,H-1}^T$ express the distance of $(H-1)^{\mathrm{th}}$ trajectory at location $(K_{k,H-1})$ till the x^{th} boundary location $K_{y,H-1}$. Afterwards obtained optical flow values are sampled in order to get efficient vector values. Therefore at optical flow process, the condensed time-axis distance A_c^u is represented in the form of feature vector as $\beta^q = [C_1^{a,q}, C, \ldots, C_{W_A}^{a,q}]$. Where β^q indicates feature vector at trajectory. α^q and $C_y^{a,q}$ denotes the time-axis space at y^{th} trajectory path for trajectory α^q.

3.3 Feature Fusion and Detection

A MHM feature is considered to identify abnormal group activity. The proposed MHM approach is efficiently able to capture motion information of the activities along with the considered OF highpoints which have distinct behavior for smaller movements. By considering all these cases, the fusion of these process leads to effective abnormality detection, The obtained features from MHP ϑ_{MHM} and OF ϑ_{OF} are fused together as,

$$\vartheta_{FSN} = \vartheta_{MHM} \cup \vartheta_{OF} \tag{19}$$

Where, ϑ_{FSN} representsthe feature integration vector which is input to bag of words (BoW) approach [33]. BoW is used to find out the behavior as normal or abnormal with respect to ground-truth report in the considered datasets.

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

4. RESULT ANALYSIS

In this section, result analysis is done using the
proposed methodology on standard data sets. The simulation is obtained using MATLAB programming language in MATLAB 2018a with system configuration; 8 GB RAM, 2GB NVIDIA Graphics card and Intel i5 processor. The UMN datasets are considered to validate the system model performance with respect to existing approaches. Moreover, the ROC (receiver operating characteristic) curve and recognition rate (R-Rate) have been utilized to measure the detection accuracy. The ROC curve is plotted in respective of TPR (true positive rate) and FPR (False positive rate), the R-Rate can be calculated by,

R-Rate=(Total number of correctly recognized images/Total number of images) *100

The dataset of UMNis considered in this study. A very popular available dataset for research purpose contains four different scenarios such as - corridor, crowd, courtyard and hit-run. The dataset is in the form of video (i.e., .avi format), where it is reformedinto frames. There are thirty number of frames at each second, but for the usage considered only two frames per second. These frames are given as input to the proposed methodology in order to detect the normal and abnormal frames using BoW classifier. There are some normal as well as abnormal frames in every scenarios. Therefore ground truth as a reference has been taken to compute performance of system model.

4.1 Recognition-Rate Computation at different considered scenarios

Figure 2(a) represents the normal crowd activity using MHM and OF. Figure 2(b)
represents abnormal crowd activity using OF and MHM of courtyard scenario. In this study different state-of-art techniques are considered such as, LMLVD (Local Mid-Level Visual-Descriptors) [35], BM (Bayesian Model) [36], MBA (Motion based approach) [36], SFM (social force model) [37], SRM (sparse reconstruction model) [38], LSM (local stat model) [39], and MDT [40]. Figure 3 shows the recognition rate at courtyard scenario, where the proposed methodology (PM) outperform 31.82% better than the SFM, 1.23% better than the LMLVD approach.

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

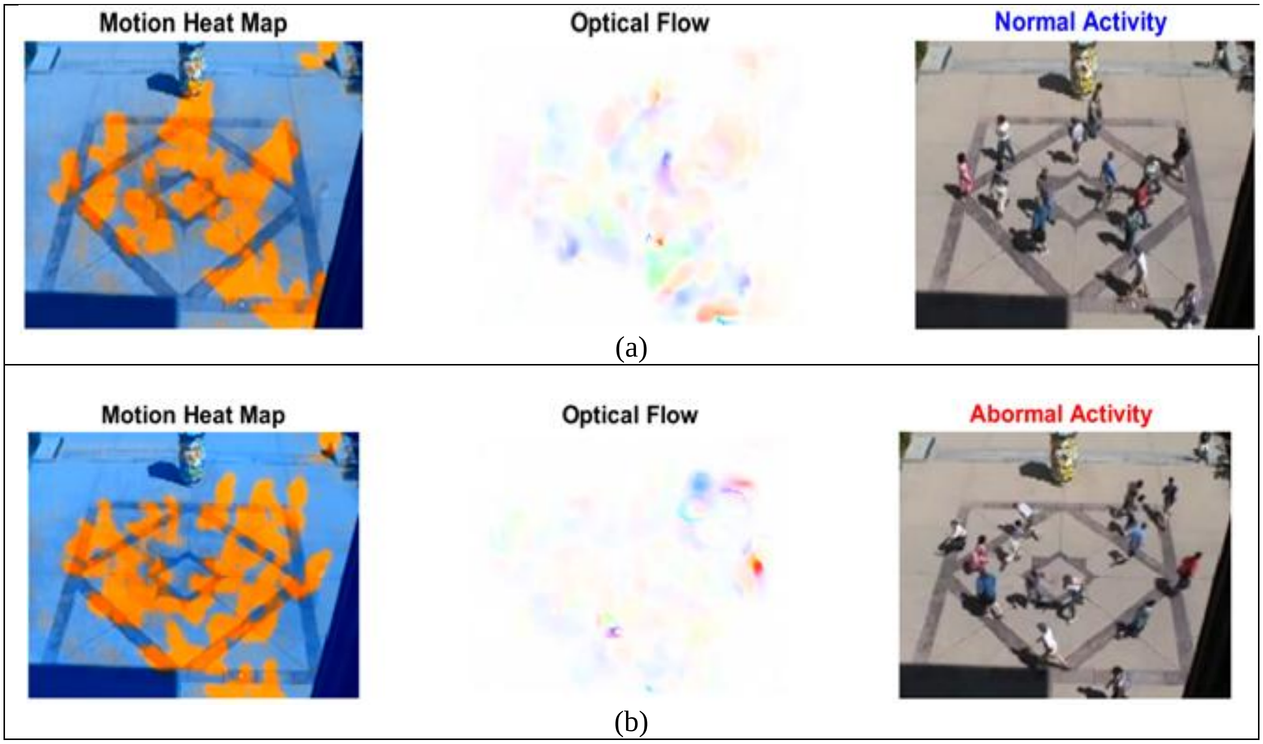

Figure:2(a) normal crowd activity using MHM and OF, (b) abnormal crowd activity using MHM and OF at courtyard

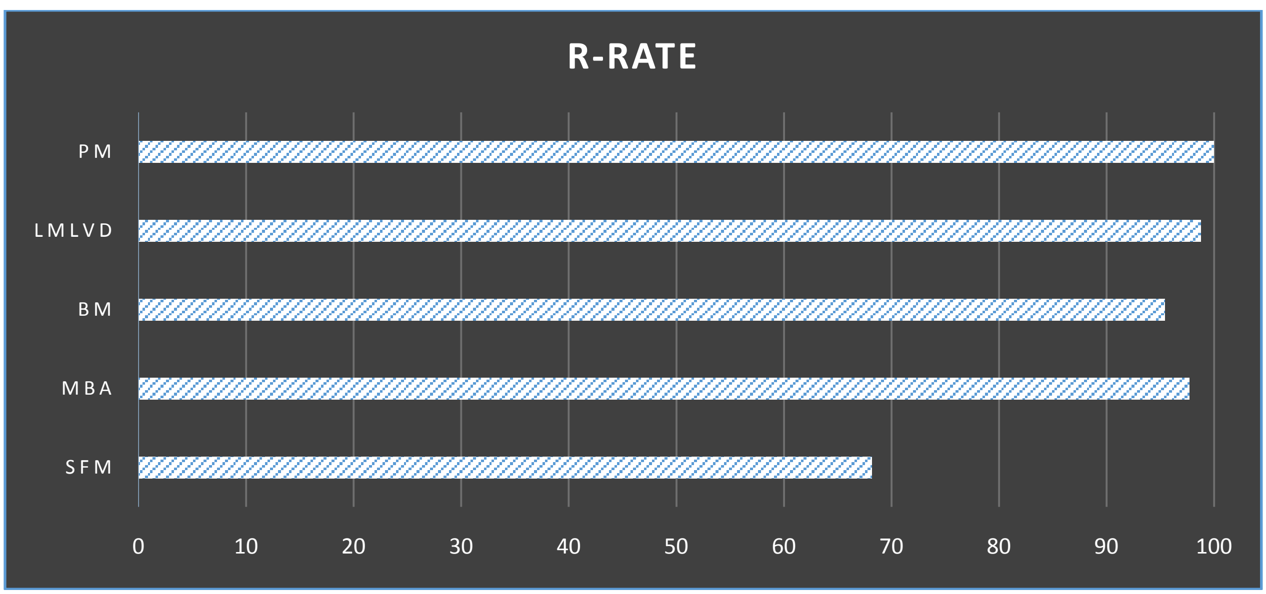

Figure 3: Recognition Rate at Courtyard scenario

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

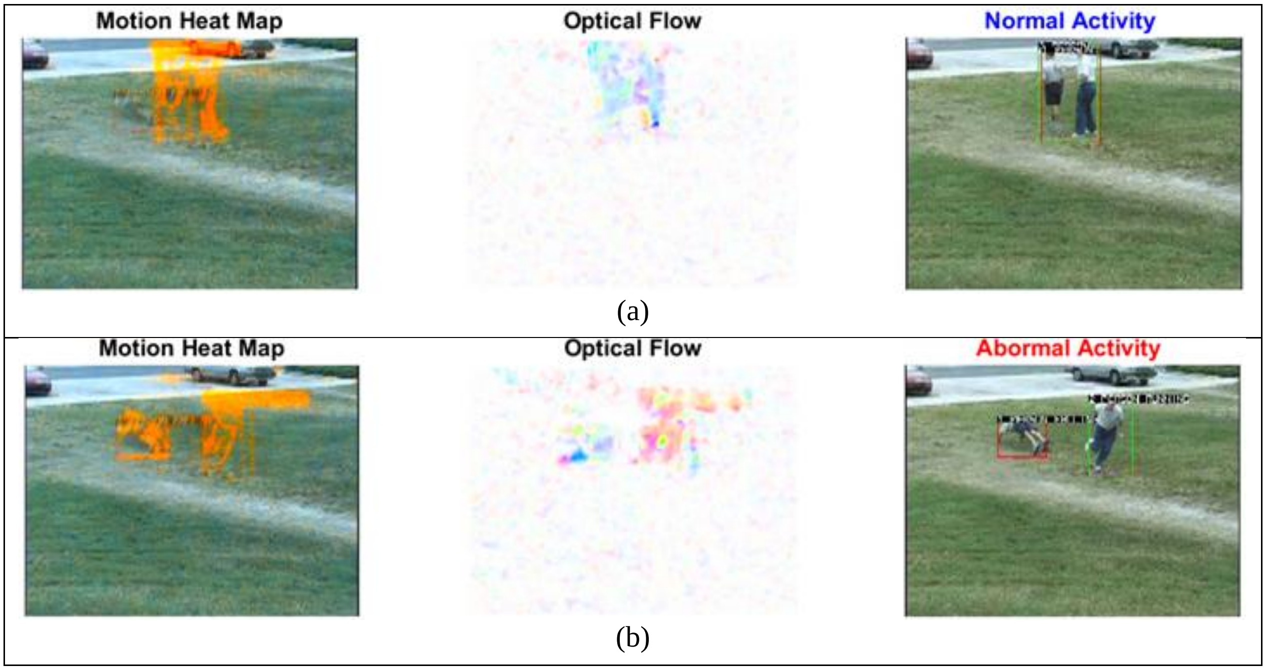

Figure 4: (a) normal crowd activity using MHM and OF (b) abnormal crowd activity using MHM and OF a hit-run scenario

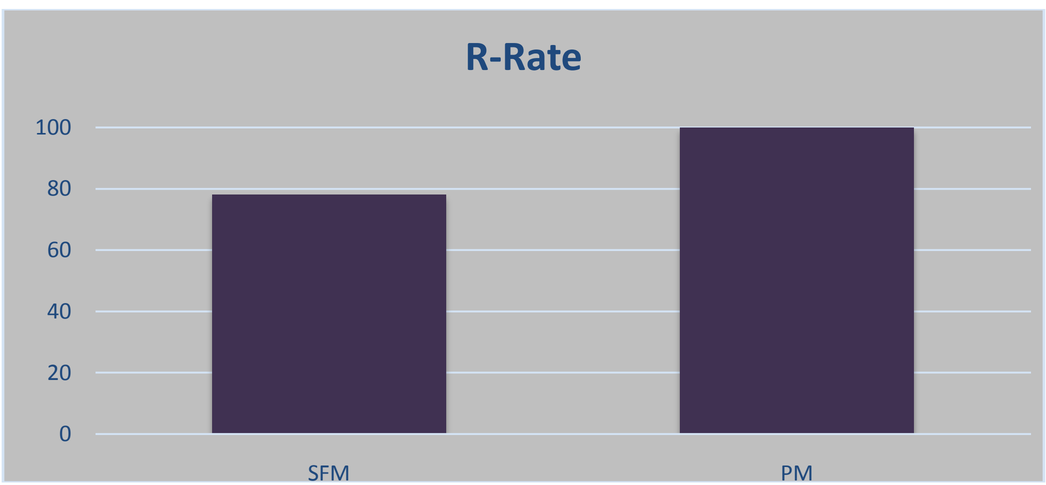

Figure 5: Recognition Rate at Hit-Run scenario

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

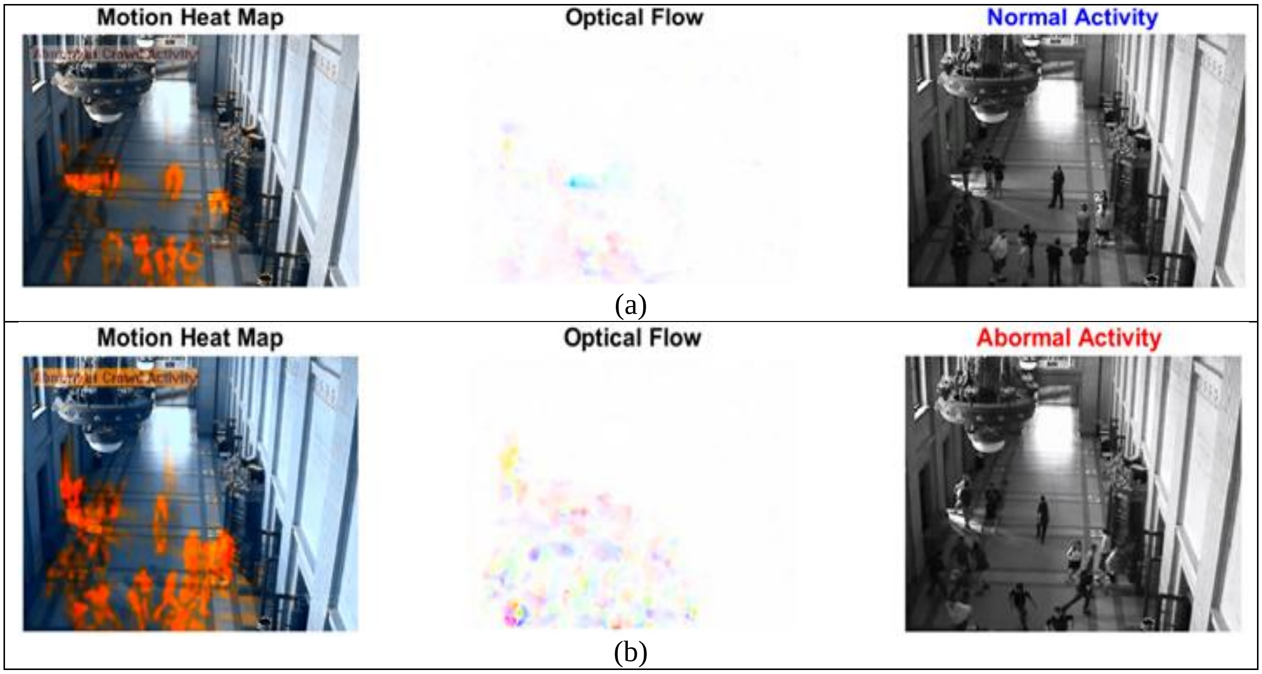

Figure 6: (a) normal crowd activity using MHM and OF (b) abnormal crowd activity using MHM and OF at corridor scenario

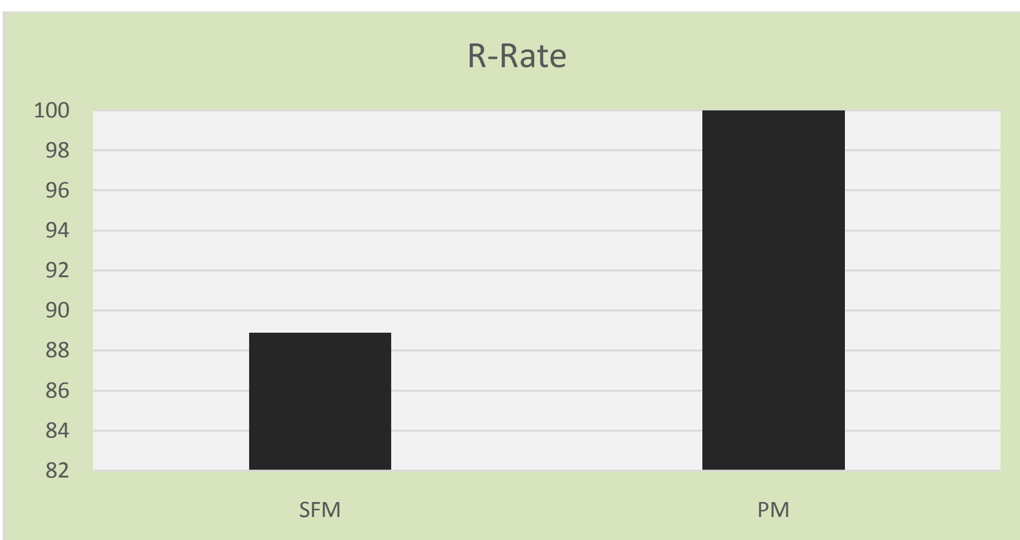

Figure 7: Recognition Rate at corridor scenario

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

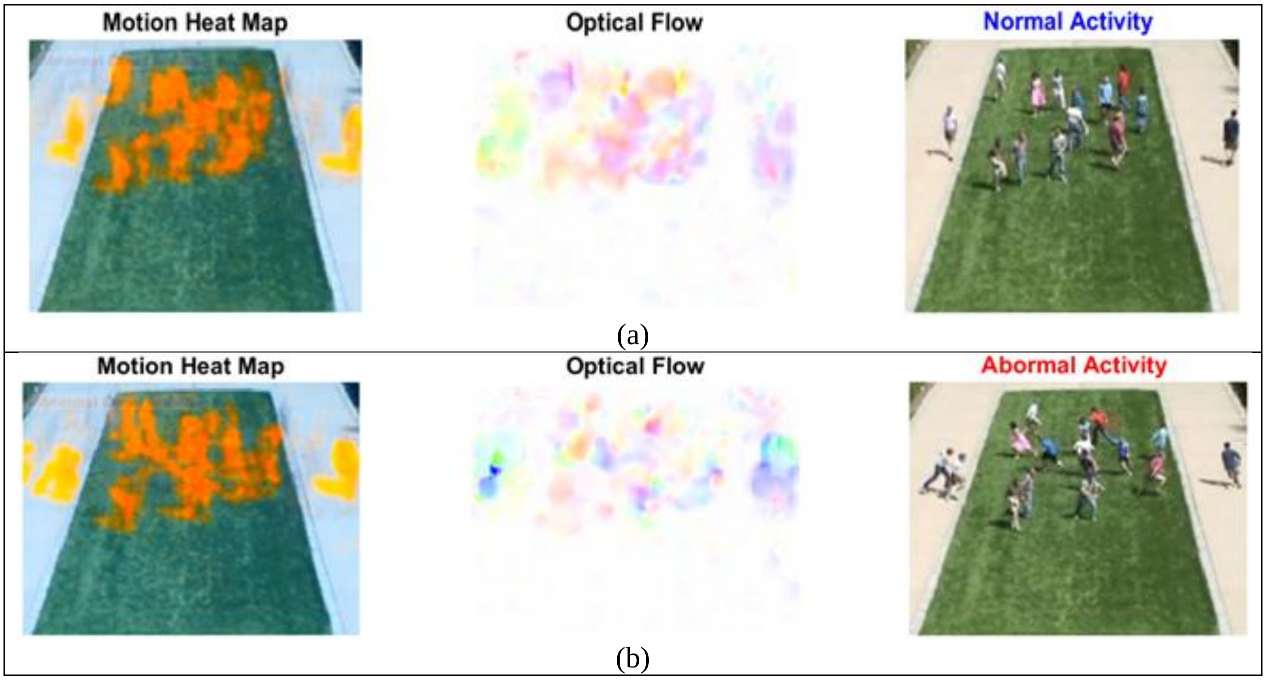

Figure 8: (a) normal crowd activity using MHM and OF, (b) abnormal crowd activity using MHM and OF at crowd scenario

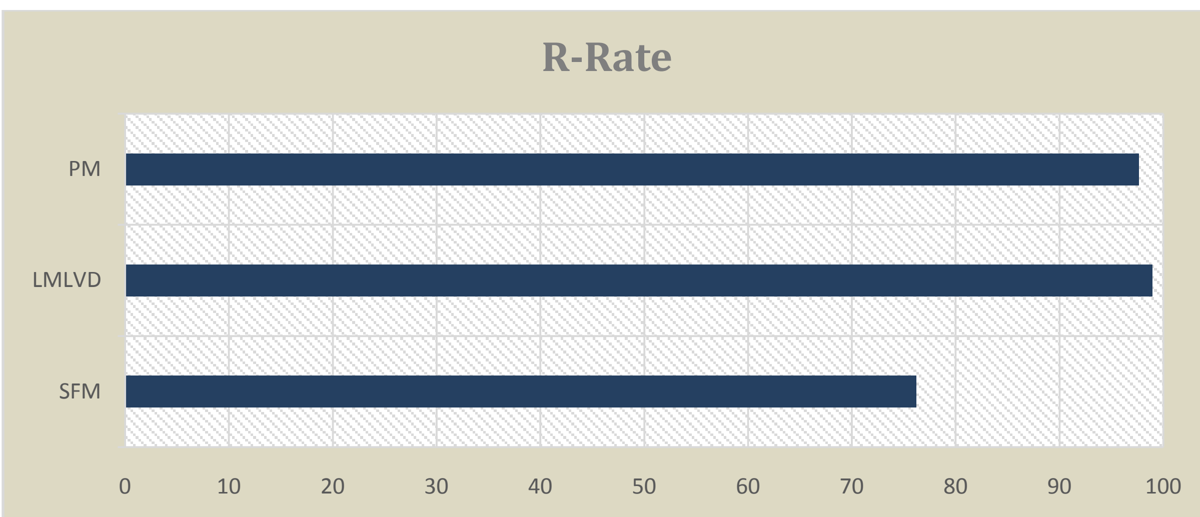

Figure 9: Recognition Rate at Crowd scenario

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

Figure 4(a) represents the normal crowd activity using MHM and OF.Figure 4(b) represents abnormal crowd activity using OF and MHM of hit-run scenario. Figure5 represents the recognition rate at hit-run scenario, where PM outperform 21.88% better than the SFM approach.Similarly, Figure 6 and Figure 8 represents obtained MHM and OF figure for corridor and crowd scenarios. Moreover, the recognition rate at corridor scenario is presented in figure 7, where PM outperform 11.11% better than the SFM approach. Correspondingly for crowd scenario, figure 9 presents the recognition rate where LMLVD have 98.94% R-Rate and SFM have 76.19% R-Rate. PM in figure 9 have 1.32% lower R-Rate as compared to LMLVD but 26.43% more R-Rate compared to SFM.

4.2 Classification using ROC and AUC comparison

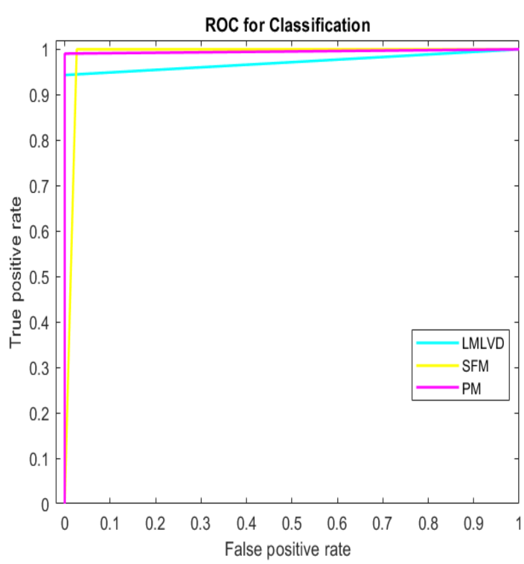

Figure 10: Classification using ROC at Courtyard scenarios

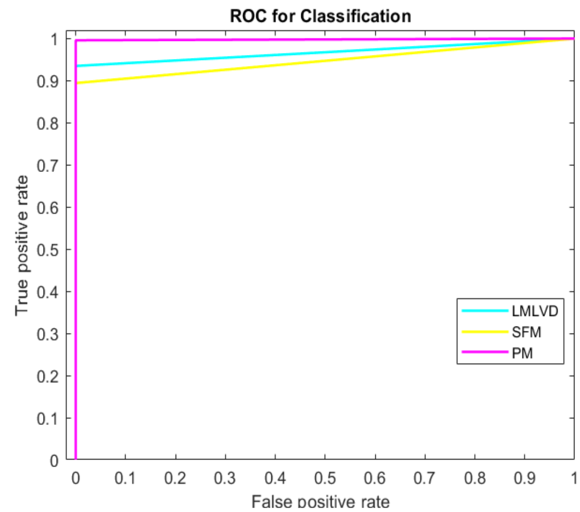

Figure 11: Classification using ROC at Corridor scenarios

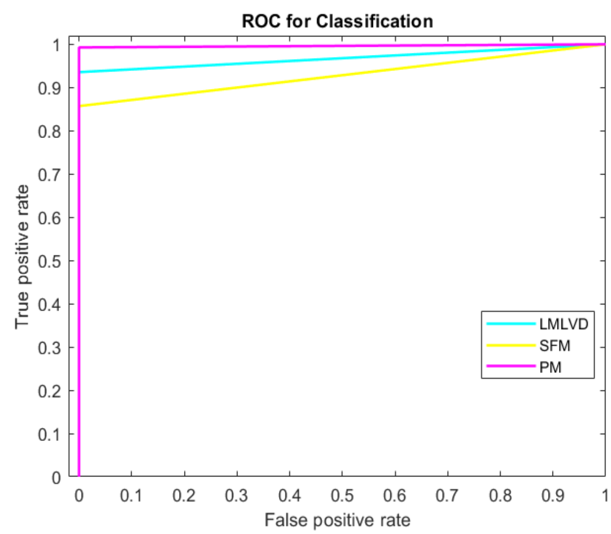

Figure 12: Classification using ROC at Crowd scenarios

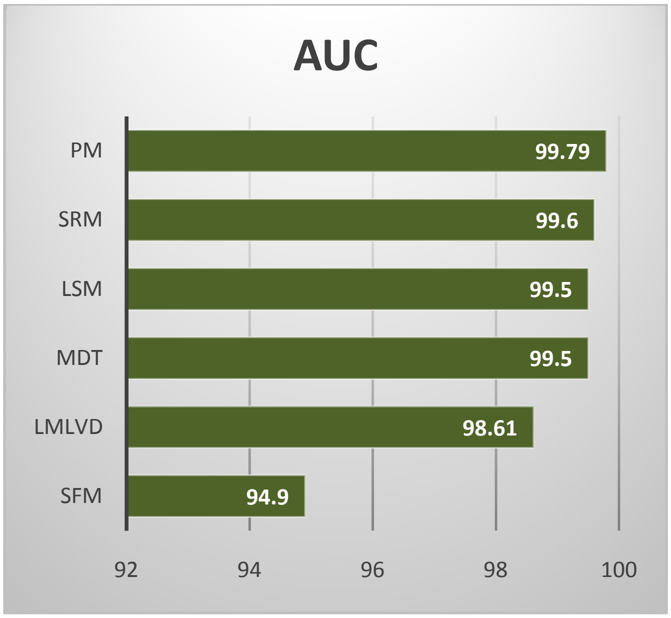

Figure 13:Over all AUC comparison with state-of-art techniques.

ROC plot has been generated for three scenarios such as; courtyard, corridor and crowd, where ROCis plotted against two existing systems LMLVD and SFM. Figure 10, 11 and 12 shows classification using ROC at courtyard, corridor and crowd scenarios. Whereas figure 13 shows over all AUC comparison with state-of-art techniques. The PM is compared with SFM, LMLVD, MDT, LSM, and SRM and we got 4.9%, 1.18%, 0.29%, 0.29% and 0.19% more AUC.

5. CONCLUSION

In this paper, decay factor has been used on some heat source so the previous heat source will become lesser important as per the greater thermal energies and new heat sources. However, the heat source thermal values can be arranged in accordance to trajectories direction and temporal information. In order to minimize the fluctuation during motion, consider the distributed process of thermal energy to diffuse originated heats by heat source series at all scenes, where features of MHM provides the motion activity information. However, it is also noted for HM that it is lacking for trajectories, therefore the trajectories reliability become very difficult to achieve. So it is essential to put on motion process of low-level to produce features as recognition process followed by optical flow. Here, fusion approach is considered by taking feature vector from MHM and OF, where integrated features is given to BoW approach to classify the frame activity. The proposed method is validated using R-Rate, ROC curve and AUC with respect to state-of-art techniques, and shows significant improvement in every considered scenarios.

6. REFERENCES

[1] J. Gantz and D. Reinsel, “The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east,” IDC iView: IDC Analyze the future, vol. 2007, pp. 1–16, 2012.

[2] Chandola V, Banerjee A, Kumar V, " Anomaly detection: A survey". ACM Computing Surveys, 2009, 41(3): 15:1-15:58.

[3] Vishwakarma S, Agrawal A,"A survey on activity recognition and behavior understanding in video surveillance". The Visual Computer, 2013, 29(10): 983-1009.

[4] C. Tu, P. Chen and N. Wang, "Improving Prediction Efficacy Through Abnormality Detection and Data Preprocessing," in IEEE Access, vol. 7, pp. 103794-103805, 2019, doi: 10.1109/ACCESS.2019.2930257.

[5] W. Ge, R. Collins, and B. Ruback, “Automatically detecting the small group structure of a crowd”, in WACV, 2009.

[6] Q. Cui, X. Liu, M. Gaoand D. Metaxas, "Abnormal detection using interaction energy potentials," in CVPR, 2011.
[7] B. Mehran, R. Moore and M. Shah, "A streakline representation of flow in crowded scenes," in ECCV, 2010.
[8] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea, "Machine recognition of human activities: A survey," IEEE Trans. on Circuits and Sys. for Video Tech., vol. 18, no. 11, pp. 1473–1488, 2008.
[9] T. Xiang and S. Gong, "Video behavior profiling for anomaly detection," IEEE Trans. on Pattern Analysis and Machine Intell., vol. 30, no. 5, pp. 893–908, 2008.
[10] F. Jiang, J. Yuan, S. A. Tsaftaris, and A. K. Katsaggelos, "Anomalous video event detection using spatiotemporal context," Comp. Vis. and Image Understanding, vol. 115, no. 3, pp. 323–333, 2011.
[11] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in IEEE Comp. Vis. and Pattern Recog. (CVPR). IEEE, 2009, pp. 935–942.
[12] X. Cui, Q. Liu, M. Gao and D. N. Metaxas, "Abnormal detection using interaction energy potentials," in IEEE Conf. on Comp. Vis. and Pattern Recog. (CVPR). IEEE, 2011, pp. 3161–3167.
[13] I. Laptev, M. Marszalek, C. Schmid and B. Rozenfeld, "Learning realistic human actions from movies," in IEEE Comp. Vis. and Pattern Recog. (CVPR). IEEE, 2008, pp. 1–8.
[14] M. Mancas, N. Riche, J. Leroyand B. Gosselin, "Abnormal motion selection in crowds using bottom-up saliency," in IEEE Image Process. (ICIP). IEEE, 2011, pp. 229–232.
[15] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, "Anomaly detection in crowded scenes," in IEEE Int. Conf. on Comp. Vis. and Pattern Recong. (CVPR), vol. 249, 2010, p. 250.
[16] W. Li, V. Mahadevan and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," IEEE Trans. on Pattern Analysis and Machine Intell., vol. 36, no. 1, pp. 18–32, 2014.
[17] T. Wang and H. Snoussi, "Detection of abnormal visual events via global optical flow orientation histogram," IEEE Transactions on Information Forensics and Security, vol. 9, no. 6, pp. 988–998, 2014.
[18] L. Kratz and K. Nishino, "Tracking pedestrians using local spatiotemporal motion patterns in extremely crowded scenes," IEEE Trans. on Pattern Analysis and Machine Intell., vol. 34, no. 5, pp. 987–1002, 2012.
[19] B. Wang, M. Ye, X. Li, F. Zhao, and J. Ding, "Abnormal crowd behavior detection using high-frequency and spatio-temporal features," Machine Vision and Applications, vol. 23, no. 3, pp. 501–511, 2012.
[20] Y.Cong,J.YuanandY.Tang,"Video anomaly search in crowded scenes via spatio-temporal motion context," IEEE Transactions on Information Forensics and Security, vol. 8, no. 10, pp. 1590–1599, 2013.
[21] H. Mousavi, M. Nabi, H. K. Galoogahi, A. Perina, and V. Murino, "Abnormality detection with improved histogram of oriented tracklets," in International Conference on Image Analysis and Processing. Springer, 2015, pp. 722–732.
[22] H. Mousavi, S. Mohammadi, A. Perina, R. Chellali, and V. Murino, "Analyzing tracklets for the detection of abnormal crowd behavior," in Applications of Computer Vision (WACV), IEEE, 2015, pp. 148–155.
[23] C. Lu, J. Shi and J. Jia, "Abnormal event detection at 150 fps in matlab," in IEEE International Conference on Computer Vision (ICCV), 2013, pp. 2720–2727.
[24] B. Zhao, L. Fei-Fei and E. P. Xing, "Online detection of unusual events in videos via dynamic sparse coding," in IEEE Computer Vision and Pattern Recognition (CVPR). IEEE, 2011, pp. 3313–3320.
[25] S. Pellegrini, A. Ess, K. Schindler and L. Van Gool, "You'll never walk alone: Modeling social

July – August 2020
ISSN: 0193-4120 Page No. 6320 - 6336

behavior for multi-target tracking," in IEEE Int. Comp. Vis. (ICCV). IEEE, 2009, pp. 261–268.

[26] Y. Cong, J. Yuan and J. Liu, "Sparse reconstruction cost for abnormal event detection," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2011, pp. 3449–3456.

[27] S. M. Masoudirad and J. Hadadnia, "Anomaly detection in video using two-part sparse dictionary in 170 fps," in 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA). IEEE, 2017, pp. 133–139.

[28] N. Li, X. Wu, D. Xu, H. Guo and W. Feng, "Spatio-temporal context analysis within video volumes for anomalous-event detection and localization," Neurocomputing, vol. 155, pp. 309–319, 2015.

[29] D. Xu, E. Ricci, Y. Yan, J. Song and N. Sebe, "Learning deep representations of appearance and motion for anomalous event detection," in British Machine Vision Conference. (BMVC), 2015.

[30] H. Carslaw and J. Jaeger, Conduction of Heat in Solids, Oxford University Press, 1986.

[31] D. Wolinski, S. J Guy, A.-H. Olivier, M. Lin, D. Manocha, and J. Pettr´e. Parameter estimation and comparative evaluation of crowd simulations. In Computer Graphics Forum, volume 33, pages 303–312. Wiley Online Library, 2014

[32] R. H. Pletcher, J. C. Tannehill and D. Anderson, Computational fluid mechanics and heat transfer. CRC Press, 2012.

[33] Ma, L., Chen, B., Wang, X., Zhu, Z., Wang, R., & Qiu, X. (2019). The analysis on the desired speed in social force model using a data driven approach. Physica A: Statistical Mechanics and ItsApplications.doi:10.1016/j.physa.2019.03.087.

[34] S. B. Wankhede, "Anomaly Detection using Machine Learning Techniques," 2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India, 2019, pp. 1-3, doi: 10.1109/I2CT45611.2019.9033532

[35] H. Fradi; B. Luvison; Q. C. Pham, "Crowd Behavior Analysis Using Local Mid-Level Visual Descriptors," in IEEE Transactions on Circuits and Systems for Video Technology, vol.PP, no.99, 2017

[36] I. R. de Almeida, V. J. Cassol, N. I. Badler, S. R. Musse and C. R. Jung, "Detection of Global and Local Motion Changes in Human Crowds," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 3, pp. 603-612, March 2017.

[37] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in IEEE on Computer Vision and Pattern Recognition, CVPR, 2009, pp. 935–942.

[38] Y. Cong, J. Yuan, and J. Liu, "Sparse reconstruction cost for abnormal event detection." in Computer Vision and Pattern Recognition, CVPR, 2011, pp. 3449–3456.

[39] V. Saligrama and Z. Chen, "Video anomaly detection based on local statistical aggregates," in IEEE on Computer Vision and Pattern Recognition, CVPR, 2012, pp. 2112–2119.

[40] W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 36, no. 1, pp. 18–32, 2014.