

Available online at <http://www.mecs-press.net/ijeme>

# Anomaly Detection in Crowded Scene by Pedestrians Behaviour Extraction using Long Short Term Method: A Comprehensive Study

Anupam Dey, Fahad Mohammad, Saleque Ahmed Raiyan Sharif, A.F.M. Saifuddin Saif

*Department of Computer Science Faculty of Information Science and Technology American International University – Bangladesh*

Received: 05 August 2018; Accepted: 22 October 2018; Published: 08 January 2019

---

## Abstract

With the expansion of worldwide security concerns and a consistently expanding requirement for successful checking of open places, i.e. air terminals, railroad stations, shopping centres, crowded sports fields, army bases or smart healthcare facilities such as daily activity monitoring and fall detection in old people's homes is increasing very rapidly. The visual occlusions and ambiguities in crowded scenes, usage of suitable method and in addition the perplexing practices and scene semantics make the investigation a challenging task. This research demonstrates comprehensive and critical analysis of crowd scene involves in object detection, tracking, feature extraction and learning from visual surveillance which helps to recognize behavioural pattern. This research refers scene understanding as scene layout, i.e. finding streets, structures, side-walks, vehicles turning, person on foot intersection and scene status such as crowd congestion, split, merge etc. The significance of the proposed comprehensive review to create crowd administration procedures and help the development of the group or people, to maintain a strategic distance from the group calamities and guarantee general society security. Based on the observation of previous research in three aspects, i.e. review based on methods, frameworks and critical existing results analysis, this research propose a framework for anomaly detection in crowded scene using LSTM (long Short-Term Method). Proposed comprehensive review is expected to contribute significantly for the investigation of behavior pattern analysis in computer vision research domains.

**Index Terms:** Anomaly detection, Pedestrian detection, LSTM, Computer Vision, Pattern Recognition.

© 2019 Published by MECS Publisher. Selection and/or peer review under responsibility of the Research Association of Modern Education and Computer Science.

---

## 1. Introduction

Human behaviour prediction through video surveillance identify, perceive, or pick up intriguing occasions

\* Corresponding author. A F M Saifuddin Saif  
E-mail address: saif@aiub.edu

which relevantly might be characterized in various aspects, i.e. suspicious event, sporadic conduct, uncommon behaviour, un-usual behaviour, abnormal behaviour, anomaly addresses the comprehensive investigation to perform intellectual real-time discovery of "interesting behaviour" to the human operator [1]. The challenge is that these occasions are uncommon and happen moderately with inconsistently and in some cases with exceptionally unfortunate negative outcomes. In previous research, all of these crowded scene analysis specifically were addressed as a unique topic under various circumstance. The programmed restriction of crowd is to perceive bunches that contains individuals who are connected, instead of limiting short connections. It results extreme testing in crowded situations, for example, open squares or extensive shopping centres, where spatial nearness alone does not decide if individuals are interacting [2,3]. Displaying human behaviour is to think about the human as a gadget with countless mental states, each with its own specific control conduct and interstate progress probabilities. The conditions of the model can be progressively composed to portray both here and longer-term (passing, after, and turning) practices [4]. However, the complexity of scenes and the misleading idea of irregular practices make anomaly detection still an exceptionally difficult task. In recent time deep learning architectures is used successfully to handle different task of computer vision for example classification of image, detecting object and recognition of activities[5,6].

This research illustrates comprehensive reviews about anomaly detection in various aspects i.e. methods, frameworks, extensive experimental results. Section 2 represents short background study which includes core problems in this research. Section 3 illustrates review based on previously used methods, frameworks and comprehensive experimental results and analysis, section 4 presents observation and discussion and finally, section 5 illustrates concluding remarks.

## **2. Related Works**

Human detection is the real occasion in the proposed research investigation to discover department. Recognition of human is mostly performed by unsupervised learning which intends to develop a model, i.e. Shape Model, Structure Model. For example, format, coordinating predicated strategies makes use of layouts for demonstrating the state of the full human body or body segments [7,8]. A bunching approach that estimates little gatherings voyaging together utilizing the thought of gathering " entitativity " characterized as far as criteria from Gestalt brain research, i.e. regular destiny (same or interrelated results), closeness (in appearance or practices), nearness and designing. In an arrangement of naturally extricated person on foot directions, the potential gatherings inside a sliding time window will be recognized utilizing various levelled bunching in view of vigorous measures registered from the uproarious directions [8, 9].

A single-person anomaly commonly falls including motionlessness, unusual speed, wrong direction [10,11]. Movement estimation strategies can experience the ill effects of the opening issue. To expand the affectability of inconsistency identification, the measure of forefront items can be broke down [12].A critical procedure in a group examination is occasion or conduct acknowledgment. It can be described by customary movement. In the early work, swarm conduct examination has been endeavored in inquire about theme of the PC vision particularly in reproduction and realistic field. Checking and demonstrating the group is not such a great amount to break down ordinary group conduct. These are alluded to as strange or unusual [13]. To illustrate previous, the overall review phases are shown in Fig 1.

## **3. Research Methodology**

Overall comprehensive analysis is composed into three subsections, i.e. existing methods and analysis, existing framework and analysis and previous experimental results analysis. Each of the subsections illustrates advantages and disadvantages in various aspects for anomaly detection.

i. Existing Methods and Analysis

Social force method is a broadly utilized group show that reproduces person on foot flow in view of Newtonian mechanics. The social forces model performs well at mimicking swarms and hierarchical optimization scheme that first runs on a coarse discretization of the trajectories based on just a few splines control points, which quickly computes an approximate shape of the maximum likelihood joint trajectory [14]. Social forces model calculates the future navigation of pedestrians. However, the model can only work on single pedestrians trajectories in time. To overcome the problems a new method named is introduced long short-term method (LSTM) can predict pedestrian behavior for huge amount of time. LSTM arranges the fundamental unit in the shrouded layer of a memory square, which contains at least one memory cells and a couple of versatile, multiplicative gating units which entryway information and yield to all cells in the square. LSTM neglects to figure out how to effectively process certain ceaseless time arrangement that are not from the earlier divided into preparing subsequence with unmistakably characterized closes. LSTM comprehends complex long time slack undertakings that have never been settled by past RNN calculations [15]. Space-time model and Hidden Markov models use LSTM to detect huge number of pedestrians moving in unlike directions.

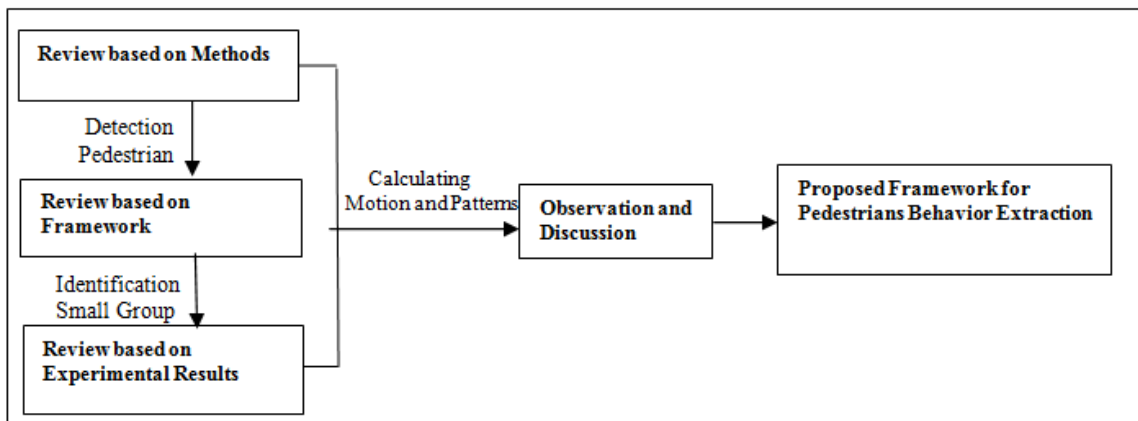


Fig.1. Framework for Analysis of Anomaly Detection In Crowded Scene by Pedestrian's Behaviour Extraction Using LSTM based Techniques.

The HMM is a grouping model with probabilistic succession and speaks to the crowd motion with an accumulation, one for each spatial area, prepared on the neighborhood spatiotemporal movement designs [16]. The paths of pedestrian are used for the construction of dynamic pedestrian graph which includes all detected people in the landscape. Analyzing the points in the graph by HMM-based, pair wise gesture between pedestrians was found and large number of people can move in unlike directions [16]. However, these models only able to visualize material goods that are not shared by the adjacent pixels which is solved by Group Behavior Model [17]. Group Behavior Model is a merge of pedestrian detector and a multi-object data association algorithm to derive long-term route of individuals passing through the landmark. The detector runs continuously and any new people reaching the landmark, people are already identified several times [17]. So solve this problem. Later, Mixture model of Dynamic pedestrian-Agents (MDA) is addressed by Zhou et al. (2012).

Mixture model of Dynamic pedestrian-Agents (MDA) considers aggregation of personal conduct standards of people on foot in swarmed scenes [18]. MDA model is unsupervised model which is designed based on genuine information by recreating group practices, gathering past practices, anticipating the future practices of

walkers given their directions just in part watched, and also characterizing distinctive person on foot practices in the scene. MDA is learned by the Shattered route of a people in the landmark which is suitable to optimal pedestrian-agent and estimate the past and future route of people; in addition the model can track down the starting and ending point of the person. [18]. However, better accuracy was previously achieved by Floor field's method introduced by Ali and Shah (2008). The accuracy of structural human behavior detection is satisfactory using Floor field method. Floor field's method has quantities of steps taken by sink-chasing procedure to achieve the sink which are known as the look for steps [19]. Floor field's method provides satisfactory accuracy in calculating structural scene, however, the method does not provide good accuracy in unstructured scene and can be found that it is more common to find unstructured human crowded Landmark than structural crowded landmark [19]. The problems of detecting unstructured crowded situation a new algorithm was developed named Correlated Topic Model (CTM) addressed by Chang et al. (2015).

Advantage of utilizing subject Models like Correlated Topic Model (CTM) is to empower to sidestep the requirement for protest location inside swarmed scenes for coordinate handling on low level stream vectors and in the meantime interface these low level highlights with abnormal state swarm practices. CTM used to catch distinctive covering and uncovering swarmed conduct in the scene in which each prospect is connected by method for an arrangement of execution greatness, where practices speak to conveyances over low-level movement highlights [20].

Table 1. Advantages and Disadvantages of Various Methods in Existing Research.

| Method Name                | Previous Research   | Advantages   | Disadvantages   |
|----------------------------|---|--|---|
| <b>Social Force Model</b>  | M. Kuderer, H. Kretzschmar, C. Sprunk and W. Burgard [14] | This investigation did in indoor and outside situations, exhibit that the approach decreases the information affiliation mistake by up to a factor of two. | The social power show makes around 30% less information affiliation mistakes  |
| <b>LSTM</b>                | A. Alahi, K. Goel, A. Robicquet, [15]                     | LSTMs can learn and imitate long arrangements, they don't catch conditions between various associated successions  | LSTM neglects to figure out how to effectively process certain ceaseless time arrangement that is not from the earlier. |
| <b>MDA</b>                 | Bolei Zhou, Xiaogang Wang and Xiaou Tang [18]             | When MDA is gained from perceptions without supervision, it can be utilized to group the directions of people on foot into varieties.                      | It cannot detect panic rising and evacuation.   |
| <b>Floor fields method</b> | Ali, S. and Shah, M. [19]                                 | It can be found that it is more common to find unstructured human crowded Landmark than structural crowded landmark.                                       | It does not work well and is inefficient in unstructured crowded scene.   |
| <b>CTM</b>                 | Rodriguez, M., Ali, S. and Kanade, T. [20]                | CTM offers an exquisite structure inside which multi-methodology of group practices and relationships among them can be taken care of                      | Correlated Topic Model (CTM) used to catch distinctive covering and uncovering swarmed conduct in the scene             |

## ii. Existing Frameworks and Analysis

In crowded scene analysis to break down segment or cluster the movements, flow direction within the region and outside is important to get some high-level interpretation and profoundly attractive scene [21]. The fundamental advantage of this framework relies on its principle focus and feature potential issues that appropriately admit the current issue and settle on what strategies to utilize (or not use) in building up a powerful identification framework for the specific setting. The impact of noisy information, the choice and portrayal of low-level features, essentially influences the discriminative intensity of the classifier. Irregular

action can be maliciously adjusted to show up as normal by the human agents [21]. In this manner, the spatiotemporal varieties for a similar movement can be high notwithstanding when performed by a similar person.

Generative graphical model methodologies are utilized to learn and perceive human activities in video as these methodologies heartily speak to inadequate spatial worldly premium focuses and an unsupervised learning approach [22]. The capacity to find significant activities naturally from the co-occurrence of visual words and utilize as a part of situations vulnerable against the unsettling influence which may bring about incorrect segmentation [22]. BNSM (Bayesian Nonparametric Scene Modelling) addresses this issue and furthermore takes care of anomaly-detection problem [23]. Since then BNSM model performs well however they are constrained to recognize particular sorts of predefined anomalies only. BNSM model framework is extremely restricted in robustness and adaptability, particularly for inconspicuous events in the scene. It must be noticed that the framework can be compared with different sorts of sensor information, image information and quantity of pixels in the image. BNSM model framework deals with the detection problem issue by utilizing Bayesian non-parametric scene understanding procedure to recognize perceptions that are confusing for other model. BNSM model framework can identify anomalous events in a non-stationary unstructured condition, for example, coral reefs. However, dimensionality of the BNSM model framework is huge; a considerable measure of preparing information is required. Consequently, if a substance is in a more dense distribution of words, it is likely going to be spoken to utilize numerous subjects under the current model [23]. By which occlusion limitation probabilistic pedestrian detection framework is proposed to detect objects and translate deformation of parts.

Considering occlusion limitation, probabilistic pedestrian detection framework is proposed to detect objects and translate deformation of parts [24]. Bottom-up hierarchical clustering approach is used to discover these small groups of pedestrians. The validation process of quantitatively and qualitatively result set based on real-world video and pedestrian scenes which is intriguing to take note that direction data alone is sufficient to yield generous concurrence with the impression of human coders who can address the gathering errand by watching more unpretentious visual signals, for example, arm motions and look heading. Through extensive experimental comparison on multiple datasets, probabilistic pedestrian detection framework's sustainability and performance is comparatively low on perspectives unlikely to occlusion [24]. However, the solution of the problem address by probabilistic pedestrian detection framework can be solved by Social-Aware Generative Adversarial imitation learning framework illustrated by Ge et al. (2009).

SA-GAIL (Social-Aware Generative Adversarial Imitation Learning) executes the limitation of probabilistic pedestrian detection framework by the underlying decision-making process of pedestrians in crowds. SA-GAIL proposes a process to distinctive factors of individual person's basic leadership is unravelled with information maximization, with the procedure demonstrated by collision avoidance regularization. Applying a novel data driven Imitation Learning system to demonstrate the consecutive basic procedure to make decision and analyse pedestrian behaviour. SA-GAIL extends the Generative Adversarial Imitation Learning framework to anticipate future paths of pedestrians by decision-making process in an unsupervised manner [25]. This research also figures out some failure cases such as de-located last observed coordinates inevitably causes large discrepancy and inability to recognize the scene such as walls.

From the above illustrated frameworks, this research observed that in some extremely crowded scenes, though similar in density, anomaly detection is not structural enough due to the high variability of pedestrian movements, hence overall detection accuracy falls down. Based on the comprehensive study, this research proposes a framework shown in Fig 2 where more attention are given on pedestrian problems and overall methodology to solve the the problem stated above in various aspects.

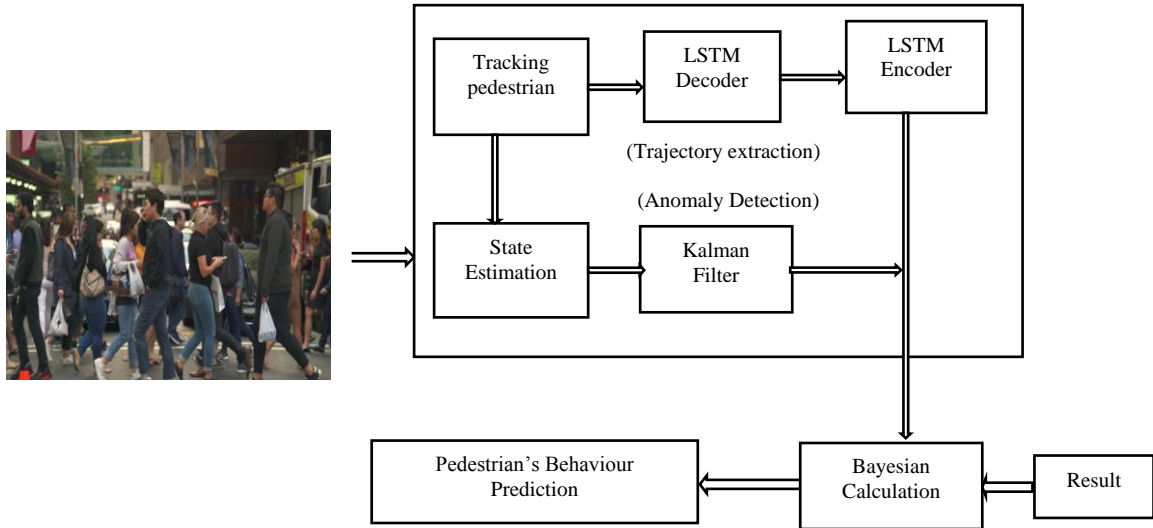


Fig.2. LSTM Based Kalman Filtering For Pedestrians Behavior Extraction.

Firstly, this research introduced microscopic level where proposed research concern about the movements of each individual pedestrian and analyze according to the collective information of them. The step is called Tracking Pedestrian .In the next step, proposed framework intends to use LSTM Decoder which includes hidden state of decoder, context vector, decoder function and decoder output result at the estimated time. Next the process will produce encoded sequence by LSTM Encoder. From the result of tracked pedestrian propose research intend to analyze and estimate the current state of the pedestrian by the mean of feature extraction. Specially, detection and understanding anomalies in crowd proposed framework aims to use Kalman recursions which are employed to optimally estimate the state sequence of anomalies and propagate an error covariance for each state. Finally, proposed framework aims to use Bayesian Calculation for the Predicted result of Pedestrian's Behaviour for accuracy estimation more efficiently.

### *iii. Experimental Results and Analysis*

This research investigates freely accessible dataset of ordinary and strange group recordings to track pedestrian, dissect and assessed the present condition of the pedestrian. From the above clarified and evaluated techniques and frameworks, this research intends to focus on uncommonly distinguishing and understanding abnormalities in crowd. Comprehensive reviews prefer to utilize Mean Deviation shown in Fig 3 on pedestrian agent record which is perfect with the Cohen's Kappa Test shown in Table 2 on the Indoor Sequences to test pedestrian directions. For more fitting assessment proposed comprehensive study prefers to analyze Kappa test alongside Adjusted Rand Index as standard factual measure. Execution of measurements is displayed as Quantitative results of the considerable number of strategies shown in Table 3. Error reduction is assessed as fundamentally and similarly in the dataset. Execution of social force model is estimated for dynamic characterization. Here we discuss some of the important methods and framework based on experimental result set.

Table 2. Cohen’s Kappa Test on the Indoor Sequences [23]

|             | SU1        |      | SU2        |      |
|-------------|------------|------|------------|------|
|             | Match rate | K    | Match rate | K    |
| dichotomous | 89%        | 0.75 | 84%        | 0.74 |
| dichotomous | 85%        | 0.69 | 75%        | 0.63 |

Kappa test is a more fitting assessment metric than understanding rate shown in Table 2. The accord ground truth and the PC assessed was a generous assertion ( $> 0:6$ ) between  $z$  for the SU1 grouping.

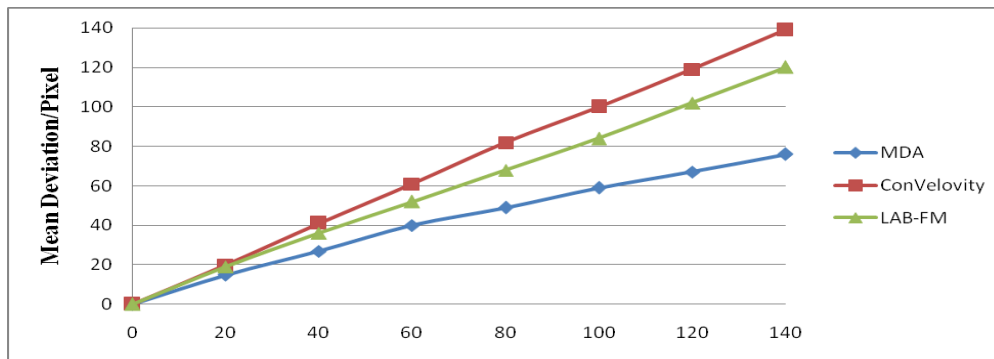


Fig.3. The Averaged Prediction Errors with Different Methods Tested on 30 Trajectories.

MDA can anticipate person on foot's practices given that their directions are just somewhat watched. The creator physically marked 30 directions (shown in Fig 3) of people on foot as ground-truth. For each ground-truth direction, the researcher utilizes the perceptions of the initial 20 edges to assess person on foot specialist list  $z$  with the calculation in Table 1 [32].

Table 3. Quantative Results among Various Methods on Various Datasets. Results are Presented based on Performance Metrics as Follows: the Average Displacement Error, rows [26].

| Metric   | Methods    | Lin  | LTA  | SF [12] | IGP* [81] | SF-mc | LSTM | O-LSTM | Social-LSTM |
|----------|------------|------|------|---------|-----------|-------|------|--------|-------------|
| Avg disp | eth [10]   | 0.80 | 0.54 | 0.41    | 0.20      | 0.41  | 0.60 | 0.49   | 0.50        |
|          | Hotel [10] | 0.39 | 0.38 | 0.25    | 0.24      | 0.24  | 0.15 | 0.09   | 0.11        |
|          | Zara [11]  | 0.47 | 0.37 | 0.40    | 0.39      | 0.35  | 0.43 | 0.22   | 0.22        |
| error    | UCY [11]   | 0.57 | 0.51 | 0.48    | 0.61      | 0.45  | 0.52 | 0.35   | 0.27        |
|          | Average    | 0.53 | 0.44 | 0.39    | 0.37      | 0.37  | 0.44 | 0.28   | 0.27        |

Social pooling based LSTM and O-LSTM overcome intensely designed Social Force and IGP models in all datasets. Specifically, the mistake diminishment is more critical on account of the UCY datasets when contrasted with ETH [26].

Table 4. Results Correlation Utilizing GDSR.

| <b>Dataset</b>    | <b>GDet</b> | <b>GTrack</b> | <b>SFM-Det</b> | <b>SFM-TR</b> | <b>JIST</b> | <b>DPMM</b> |
|-------------------|-------------|---------------|----------------|---------------|-------------|-------------|
| <b>ETH</b>        | 78 %        | 80 %          | 77 %           | 78 %          | 54 %        | 63 %        |
| <b>HOTEL</b>      | 89 %        | 89 %          | 78 %           | 81 %          | -           | -           |
| <b>Student003</b> | 71 %        | 72 %          | 58 %           | 60 %          | -           | -           |

In table 4, GDet refers proposed aggregate discovery, GTrack refers proposed bunch following on the yield of GDet, SFM-Det refers gathering recognition, SFM-TR refers proposed assemble following on the yield of SFM-Det [34]. The execution of GDet and GTrack is more suitable comparing with various methods discussed above. HSFT has a superior execution contrasted and the unadulterated optical stream, coordinate cuboids component extraction and wavelet change in the (x, y) plane [27]. HSDT has better demonstration which is proposed for dynamic portrayal of irregular conduct identification.

#### 4. Observation and Discussion

This research illustrates observation into three aspects, i.e. observation on current burdens, observation on methods and observation on datasets for anomaly detection. The next subsequent section represents current challenges followed by potential methods which can be addressed to overcome the challenges. After that, based on comprehensive analysis on previous datasets, observation on various datasets is presented for experimental validation.

##### *i. Observation on current burdens on anomaly detection*

Automated pedestrians recognition can collect directions from video and that progressive bunching can distinguish little gatherings of individuals voyaging together and Optical streams are first assessed and after that utilized for a piece of information to group human group into bunches in unsupervised way utilizing the proposed AMC algorithm by the comprehensive review [28,29]. The burdens originate from the diverse trademark as posture, position and thickness, changes of movement over the long haul in a genuine video succession, light changes and various information channel with the measure of camera consists of numerous sides and Mixture model of Dynamic Pedestrian-Agents to take in the aggregate elements from video in swarmed scenes [30, 31].

##### *ii. Observation on methods on anomaly detection*

Discriminative graphical model can be at the same time misuse intra-camera and between camera settings for movement identification without topology derivation and to distinguish various group practices through dependability examination for dynamical frameworks, without the requirement for question recognition [32, 33]. Visual practices are mind boggling and have much assortment in an unconstrained domain and the impact of boisterous information, the decision and portrayal of low-level highlights, altogether impacts the discriminative intensity of the classifier and conduct model can adjust to client characterized situation and create reasonable human practices in a group [34,35].

By testing the viability of HFST on unusual location LDA is utilized and enhanced numerous HMMs to display the worldwide and neighborhood scenes, separately. This approach performs superbly in both productivity and discovery exactness and CTM refers to every scene which is related with an arrangement of conduct extents, where practices speak to conveyances over low-level movement highlights [36, 37]. Human conduct is simply the result of connecting determinants, for example, self-inspiration, social impacts, and



ecological occasions. Express social impacts are effects on client conduct achievement from coordinate associations in a sound interpersonal organization. This is inadequate to completely portray social impacts [38].

The perception stage finds the area of the picture that is more like the objective, accepting as earlier the speculation is given in the dynamical stage. Following, and particularly the dynamic module, is profited by Social Signal Processing method. Such module does not consider that individuals, at whatever point allowed moving in a substantial domain [39]. During experimentation in genuine situations, such module observed high fluctuation in human movement and it filled in as the inspiration to build up the present work. Given the high fluctuation of human conduct, it is hard to precisely anticipate human movement utilizing a similar arrangement of parameters, since high fluctuation of human conduct shows up a reasonable distinction amongst expectations and perceptions [40].

LSTM, Hidden Markov and social force models does not work well in large and complex crowded scene. Social field model depends on small groups in the crowd but the performance of the model in complex scenario is not up to the desired thrust. Hidden Markov Model methodologies are normally constrained to a little, known number of people due to the combinatory associated with the coupled elucidation of different time arrangement. Group Behavior model depends on agglomerative hierarchical clustering. For this reason, the calculation time in this process for detecting a small group takes up to four minutes. In the scenario of behavior detection it might take up to 10- 15 minutes which is computationally expensive and is not efficient enough. Group Behavior model depends on neural network with huge trained data for more accuracy and instant results. Flood field method cannot deal with unstructured crowded scenario. CTM (Correlated Topic Model) and MDA (Mixture Model of Dynamic Pedestrian-Agents) deal with structural and unstructured behavior. Problem with MDA is the need of starting point and ending point to be initialized and both the model worked with 2 to 4 behavior. There are about 8 basic Questionable behaviors which have been ignored in the respective experimental results. These models are not that dynamic, some parameters are required to initialize for getting a desired output. CRF and MDT methods to oddity location with different restriction utilizing camera by checking the camera design. Thus, the size of estimation is being reliant on number of human distinguished from video or pictures [41]. At the same time recognition and following of human in swarmed scene turn out to be extremely troublesome and location of inconsistency in view of surroundings in addition relies upon spots, network and event. In the scene of night tracker legitimately worked with person on foot's course of development and DBN-ANN and DBN-R strategy to utilize and anticipate people on foot next appearance [42, 43]. For betterment of detecting some combination of Group behavior method, CTM (Correlated Topic Model) and mixture model of dynamic pedestrian-Agents (MDA) is recommended by this comprehensive review.

### *iii. Observation on datasets on anomaly detection*

The dataset contains different observation recordings, with extensive group. Anomalies display in the dataset include: bikers videos, engine vehicles videos, railway stations, bus stations videos, images. Cohen's Kappa Test on the Indoor Sequences to test pedestrian's bearings. Besides, MDA can envision individual by walking's practices given that their bearings are simply restricted to some degrees. In both cases, result set sufficient for some selected videos and datasets. Social forces based LSTM and O-LSTM smidgen superior to other strategy found in this comprehensive review. For all dataset, achievement is over 75% which is truly great. All of those methods do not have the ability to pick up anomalies even at the far end of the scene. Proposed review also compared and noted false positive rate and true positive rate for all of the methods of detecting anomaly from crowded scene for better investigation the utilization of more descriptive features, the introduction of movement which would permit the detection of events, for example, wrong-way movement of pedestrians. From the discussion of those methods and frameworks based on experiment, the result set has some major problems to appear on anomaly detection and analysing. Datasets collection should be more perfect for analysing properly. More research and analysis need to perform to increase detection rate and take decision on those events instantly.

Proposed comprehensive review concerns about the developments of every individual pedestrian at the tiny level and investigate as indicated by the aggregate data of them. In Some situation the degree of crowded scenes is extreme, however comparable in density, are less basic because of the high fluctuation of pedestrian movements. Since pedestrian detection is a standout amongst the most critical themes in object detection, it has pulled in much consideration as of late. In this paper, overall review is construed from the presence of single pictures through investigating the solid connections among the visibilities of various parts with numerous sizes. The principle issue is to set as a regulated characterization issue of unprecedented practices. Consequently, if a substance is in a more dense distribution of words, it is likely going to be addressed using various subjects which are the discernible impediment of Bayesian Nonparametric Scene Model. Through test correlation on different datasets, its maintainability and execution is similarly low on perspectives unlikely to occlusion, so if there is no occlusion happens the proposed methodology is not expected to deliver desired experimental results. In addition, failing to de-locate last observed coordinates inevitably makes substantial disparity and inability to recognize the scene such as walls.

## 5. Conclusion

Recently, interest is increasing in computational methodologies for pretending and examining the activities and behaviours of crowds from real world videos which makes behavioural pattern analysis more fascination for further investigation. This research illustrated comprehensive review of the state-of-the-art methodologies for crowded scene analysis across some key aspects. Based on the reviews based on various aspects, this research proposes a framework for tracking, state estimation and predicting pedestrian's future behaviour. Later, this research presents a shared approach for figuring trajectory level behaviour features from crowd videos which is expected to handle normal and moderately dense crowd videos, and also can calculate the trajectory and movement behaviour for each pedestrian' sin each time step and has a major positive impact when supervision longer trajectories in heavily disorderly localities. One of the key advantages of the proposed approach is expected to detect changing movement significantly and behaviours of pedestrians and finally, estimation. Although a variety of delegacy approach path and modelling have been proposed and a lot of works has been done, numerous issues in crowded scene analysis are as yet open, and they deserve further research which can be further addressed and addressed to solve using the comprehensive review demonstrated by this research.

## References

- [1] T. Boult, "Frame-rate multi-body tracking for surveillance," in Proc. DARPA Image Understanding Workshop, Monterey, CA, Nov. 1998, pp. 305–308.
- [2] L. Bazzani, M. Cristani, and V. Murino. Decentralized particle filter for joint individual-group tracking. In Proc. of IEEE CVPR, 2012.
- [3] M. Zanotto, L. Bazzani, M. Cristani, and V. Murino. Online Bayesian nonparametrics for group detection. In Proc. of BMVC, 2012.
- [4] Boer, E., Fernandez, M., Pentland, A., & Liu, A. (1996). Method for evaluating human and simulated drivers in real traffic situations. In IEEE Vehicular Tech. Conf. (pp. 1810–1814). Atlanta, GA.
- [5] A. Krizhevsky, I. Sutskever, and G.E. Hinton. Imagenet classification with deep convolutional neural networks. In NIPS, 2012.
- [6] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In CVPR, 2014.
- [7] D. Nguyen, W. Li and P. Ogunbona, "Human detection from images and videos: A survey", Pattern Recognition, vol. 51, pp. 148-175, 2016.

- [8] O. Popoola and Kejun Wang, "Video-Based Abnormal Human Behavior Recognition—A Review", *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 865-878, 2012.
- [9] WeinaGe, R. Collins and R. Ruback, "Vision-Based Analysis of Small Groups in Pedestrian Crowds", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012.
- [10] D. Chen and P. Huang, "Motion-based unusual event detection in human crowds", *Journal of Visual Communication and Image Representation*, vol. 22, no. 2, pp. 178-186, 2011.
- [11] V. Reddy, C. Sanderson and B. Lovell, "Improved anomaly detection in crowded scenes via cell-based analysis of foreground speed, size and texture", *CVPR 2011 WORKSHOPS*, 2011.
- [12] G. Mariem, E. Ridha and Z. Mourad, "Detection of Abnormal Movements of a Crowd in a Video Scene", *International Journal of Computer Theory and Engineering*, vol. 8, no. 5, pp. 398-402, 2016.
- [13] Y. Kim and T. Moon, "Human Detection and Activity Classification Based on Micro-Doppler Signatures Using Deep Convolutional Neural Networks", *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 8-12, 2016.
- [14] M. Kuderer, H. Kretzschmar, C. Sprunk and W. Burgard, "Feature-Based Prediction of Trajectories for Socially Compliant Navigation", *Robotics: Science and Systems VIII*, 2012.
- [15] A. Alahi, V. Ramanathan, K. Goel, A. Robicquet, A. Sadeghian, L. Fei-Fei and S. Savarese, "Learning to Predict Human Behavior in Crowded Scenes", *Group and Crowd Behavior for Computer Vision*, pp. 183-207, 2017.
- [16] Kratz, L. and Nishino, K. (2012). Tracking Pedestrians Using Local Spatio-Temporal Motion Patterns in Extremely Crowded Scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5), pp.987-1002.
- [17] WeinaGe, Collins, R. and Ruback, R. (2012). Vision-Based Analysis of Small Groups in Pedestrian Crowds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5), pp.1003-1016.
- [18] Bolei Zhou, Xiaogang Wang and Xiaoou Tang (2012). Understanding collective crowd behaviors: Learning a Mixture model of Dynamic pedestrian-Agents. 2012 IEEE Conference on Computer Vision and Pattern Recognition.
- [19] Ali, S. and Shah, M. (2008). Floor Fields for Tracking in High Density Crowd Scenes. *Lecture Notes in Computer Science*, pp.1-14.
- [20] Rodriguez, M., Ali, S. and Kanade, T. (2009). Tracking in unstructured crowded scenes. 2009 IEEE 12th International Conference on Computer Vision.
- [21] Li, T., Chang, H., Wang, M., Ni, B., Hong, R. and Yan, S. (2015). Crowded Scene Analysis: A Survey. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(3), pp.367-386.
- [22] Popoola, O. and Kejun Wang (2012). Video-Based Abnormal Human Behavior Recognition—A Review. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6), pp.865-878.
- [23] Girdhar, Y., Walter Cho, Campbell, M., Pineda, J., Clarke, E. and Singh, H. (2016). Anomaly detection in unstructured environments using Bayesian nonparametric scene modeling. 2016 IEEE International Conference on Robotics and Automation (ICRA).
- [24] Wanli Ouyang and Xiaogang Wang (2012). A discriminative deep model for pedestrian detection with occlusion handling. 2012 IEEE Conference on Computer Vision and Pattern Recognition.
- [25] W. Ge, R. Collins and B. Ruback, "Automatically detecting the small group structure of a crowd", 2009 Workshop on Applications of Computer Vision (WACV), 2009.
- [26] Alahi, A., Ramanathan, V., Goel, K., Robicquet, A., Sadeghian, A., Fei-Fei, L. and Savarese, S. (2017). Learning to Predict Human Behavior in Crowded Scenes. *Group and Crowd Behavior for Computer Vision*, pp.183-207.
- [27] Mazzon, R., Poiesi, F. and Cavallaro, A. (2013). Detection and tracking of groups in crowd. 2013 10th IEEE International Conference on Advanced Video and Signal Based Surveillance.

- [28] WeinaGe, R. Collins and R. Ruback, "Vision-Based Analysis of Small Groups in Pedestrian Crowds", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012.
- [29] D. Chen and P. Huang, "Motion-based unusual event detection in human crowds", *Journal of Visual Communication and Image Representation*, vol. 22, no. 2, pp. 178-186, 2011.
- [30] G. Mariem, E. Ridha and Z. Mourad, "Detection of Abnormal Movements of a Crowd in a Video Scene", *International Journal of Computer Theory and Engineering*, vol. 8, no. 5, pp. 398-402, 2016.
- [31] B. Zhou, X. Tang and X. Wang, "Learning Collective Crowd Behaviors with Dynamic Pedestrian-Agents", *International Journal of Computer Vision*, vol. 111, no. 1, pp. 50-68, 2014.
- [32] Z. Zha, H. Zhang, M. Wang, H. Luan and T. Chua, "Detecting Group Activities With Multi-Camera Context", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 5, pp. 856-869, 2013.
- [33] B. Solmaz, B. Moore and M. Shah, "Identifying Behaviors in Crowd Scenes Using Stability Analysis for Dynamical Systems", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 2064-2070, 2012.
- [34] O. Popoola and Kejun Wang, "Video-Based Abnormal Human Behavior Recognition—A Review", *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 6, pp. 865-878, 2012.
- [35] L. Luo, S. Zhou, W. Cai, M. Low, F. Tian, Y. Wang, X. Xiao and D. Chen, "Agent-based human behavior modeling for crowd simulation", *Computer Animation and Virtual Worlds*, vol. 19, no. 3-4, pp. 271-281, 2008.
- [36] 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, 2011.
- [37] M. Rodriguez, S. Ali and T. Kanade, "Tracking in Unstructured Crowded Scenes", *IEEE Transactions on Pattern Analysis*, vol. 34, no. 5, pp. 987-1002, 2009.
- [38] N. Phan, D. Dou, B. Piniewski and D. Kil, "A deep learning approach for human behavior prediction with explanations in health social networks: social restricted Boltzmann machine (SRBM+)", *Social Network Analysis and Mining*, vol. 6, no. 1, 2016.
- [39] M. Cristani, R. Raghavendra, A. Del Bue and V. Murino, "Human behavior analysis in video surveillance: A Social Signal Processing perspective", *Neurocomputing*, vol. 100, pp. 86-97, 2013.
- [40] G. Ferrer and A. Sanfeliu, "Behavior estimation for a complete framework for human motion prediction in crowded environments", *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014.
- [41] S. Amraee, A. Vafaei, K. Jamshidi and P. Adibi, "Anomaly detection and localization in crowded scenes using connected component analysis", *Multimedia Tools and Applications*, 2017.
- [42] V. Reddy, C. S and B. C. Lovell, "Improved anomaly detection in crowded scenes via cell-based analysis of foreground speed, size and texture", 2011.
- [43] S. Choi, E. Kim and S. Oh, "Human Behavior Prediction for Smart Homes Using Deep Learning.", *IEEE*, 2013.

## Authors' Profiles



**Anupam Dey** is a student of undergraduate (UG) program majoring in Computer Science & Engineering from American International University-Bangladesh. He is also working as an embedded software engineer for last one year. He was awarded "Dean List Honors" in AIUB. He has contributions in open source libraries. His research interest focuses but not limited to Image Processing, Computer Vision, Augmented Reality, Virtual Reality, Artificial Intelligence, Face Tracking and Machine Learning.



**Fahad Mohammad** is an undergraduate student of Computer Science and Engineering under the Department of Science and Information Technology of American International University Bangladesh. Currently, he is doing a thesis on “Human Behaviour Prediction In Crowded Scene”. His research area focused on Computer Vision, Artificial Intelligence, Natural Language & Image Processing, and Machine Learning & Neural Networks. He is also working as an asp dot net developer in Code Novo Private Limited. He has a remarkable enthusiasm for solving the real-world problem by the latest technology existed.



**Saleque Ahmed** is an undergraduate student of Computer Science and Engineering under the Department of Science and Information Technology of American International University-Bangladesh. His research interest based on Computer Vision and Pattern Recognition, Image Processing, Object Detection and Face Tracking. He is also an expert in Telnet, FTP, TCP/IP, STP, DNS, Windows & Web Server Setup, and Maintenance.



**Raiyan Sharif** is a student of undergraduate program (BSC) majoring in Computer Science & Engineering from American International University-Bangladesh. He has good skill in problem-solving. He has completed the LICT course on machine learning and Android development. His research interest is not limited to only image processing, IOT, machine learning, Pattern recognition.



**A.F.M. Saifuddin Saif** received PhD from Faculty of Information Science and Technology, University Kebangsaan Malaysia (UKM) in 2016. He received M.Sc. in Computer System Engineering (Software System) from University of East London, UK and B.Sc. (Eng.) degree in Computer Science and Engineering from Shahjalal University of Science and Technology, Bangladesh in 2012 and 2008, respectively. Most of his contributions in Computer Vision and Artificial Intelligence Research field were published in ISI Q1 journals. He has published many papers in ISI indexed Journals, Scopus indexed Journals, Book Chapters, Conferences and Proceedings. He served as Technical Committee Members, Reviewers, Guest Speakers, Session Chairs in many Conferences and Workshops. Currently, Dr. A.F.M. Saifuddin Saif is an Assistant Professor at Faculty of Information Science and Technology, American International University – Bangladesh. Before joining the university, he did Post Doctorate at Faculty of Information Science and Technology, University Kebangsaan Malaysia. He spent more than 6 years in IT industry such as Advanced Software Development, Web eMaze etc as IT researcher. His research interests include Image Processing, Computer Vision, Artificial Intelligence, Augmented Reality, 3D reconstruction and Medical Image Processing.

**How to cite this paper:** Anupam Dey, Fahad Mohammad, Saleque Ahmed, Raiyan Sharif, A.F.M. Saifuddin Saif, "Anomaly Detection in Crowded Scene by Pedestrians Behaviour Extraction using Long Short Term Method: A Comprehensive Study", International Journal of Education and Management Engineering(IJEME), Vol.9, No.1, pp.51-63, 2019.DOI: 10.5815/ijeme.2019.01.05