# Abnormal Human Behavior Detection and Classification In Crowd Using Image Processing

[1]Megha Chhirolya and [2]Dr. Nitesh Dubey
[1] PG Student, [2]Associate Professor
[1,2]Computer Science Engineering, Global Nature Care Sanghthan group of Institutions, Jabalpur, India.

***Abstract***— Monitoring and inferring socio-cognitive behaviors through crowd analysis can help us to understand many processes. Be it people in crowded environments, road traffic or even a flock of fish, situational awareness becomes critical for creating adequate disaster response, providing incident management, exercising control, etc. Recent researches have indicated that crowd modeling is conventionally based on density analysis. However, socio-cognitive behavior studies have demonstrated that crowds often display a wide variety of behaviors that arise spontaneously from the collective motions of unconnected individuals. Therefore, behavior analysis employing physics-based approaches only, thereby neglecting the socio-psychological aspects, may present diverse challenges to accurate inference. This means that by identifying and modeling some of the interacting agents that underpin the evolution of such behaviors, we can deliver contexts that can help in the autonomous analysis of social and antisocial behaviors in crowded environments.

Finally, we proposed a method based on Optical Flow features to detect abnormal behaviors and classifications are provided toward the advancement in the field.

***Keywords:*** Abnormal Behavior, Kinetic Energy, Image frames, Crowd Behavior, Optical Flow, Classifications.

## I. INTRODUCTION

Crowd behavior analysis is an emerging research area which was mainly inspired by the security concerns surrounding crowded environments. Thus, various applications of crowd behavior analysis are tackled including human motion prediction [1], crowd instability analysis [2], interacting groups detection and tracking [3], crowd behavior simulation [4], and crowd emotion detection [5]. Recently, research work has been focused on developing autonomous crowd analysis methods, and considerable progress has been achieved [6]. Nevertheless, many solutions for problems that directly impact the visual behavior analysis of the crowd still have huge room for improvements such as accurate and robust target detection and tracking, occlusion handling, inference, and modeling.

From the perspective of computer vision, crowd modeling and analysis consists mainly in tracking and trajectory analysis, which demands human detection methods that can be highly challenging in unconstrained conditions. This implies that the crowd is defined as a number of mobile targets (individual persons or groups) that demonstrate some sort of coherent behavior. On the other hand, computational psychologists are more concerned about applying path analysis and decision-making theories to model and identify specific crowd behavior patterns in certain situations, e.g. evacuation [7]. Crowd behavior analysis can be performed through three main steps as in Fig. 1.1, which is modeling, feature extraction, and behavior analysis and identification.

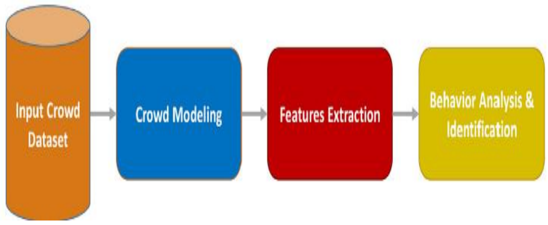



Figure 1.1: Crowd behavior analysis steps.

A core aspect that affects crowd behavior analysis is the level at which the analysis is performed. Macro level methods study the behavior of the crowd in a holistic manner, where they recognize the crowd as one entity or a number of interacting people that share the same spatial and behavioral features. At the micro-level, the crowd behavior is deduced based on the analysis of individual entities of the crowd by generating collective scene semantics out of the individual targets. Performing micro-level analysis of the crowd can be highly challenging especially in dense crowd scenes which results in high occlusion, increased appearance similarities between individuals, and added computational complexity. Consequently, some methods [8] [9] head towards the combination of both levels to provide more informative representation of the crowd that allows a deeper understanding of the crowd behavior from either perspective, and analyze the interactions between individuals and groups in the crowd to infer their impact on the behaviors.

A. Group Behavior: In group motion analysis, the socio-cognitive behavior of the crowd is studied based on the evolution of the motion of groups of spatially associated individuals. The initial essential task in such techniques is to identify and separate each group with unique characteristics and to study its motion patterns to understand the group's socio-cognitive behavior. Here, the focus is not on the characteristics of each individual within the group, since the group as one entity is assumed to have a uniform motion pattern. Such techniques are more practical in scenes with higher densities of crowds, where people moving together tend to have the same behavior patterns. In certain scenarios (where individual member of a group become detectable) individualistic behaviors can be often derived from this category.
Such group behavior can be observed in Fig. 1.2.

Figure1.2: Frames showing group behavior. [10]

In Frames above where simple groups are manually annotated with red circles and a yellow circle indicates a group with more complex behavior. In frame (a), the crowd consists of three groups, two of which are simple groups that have uniform motion patterns. However, one group is more complex since not all of its members have the same motion pattern (one person is moving in the opposite direction). Frame (b) also shows a simple group behavior where the crowd consist of four such groups.

B. Social Interaction Behavior: Socio-cognitive crowd behavior analysis can be alternatively performed by capturing the interactions of individuals or groups, with other individuals, groups, or obstacles. The need for such analysis appears neither when neither the group is uniform (as assumed in group behavior) nor the group members' act independently (as assumed in individualistic behavior). Then, the interactions can be modeled by identifying the underlying social forces at both micro and macro levels. Such analysis at different levels provides a deeper understanding to the causes and results of the socio-cognitive behavior of people in a crowd whether as individuals or groups, and how they affect one another. Social interaction behavior detection requires algorithms (normally uniquely designed for specific types of interactions) which can deduce those interactions from features representing selected relations between targets exhibiting either individualistic or group behavior [11]. Therefore, this is the most diversified category, often requiring sophisticated approaches to identify certain interactions from visually detectable clues. Nevertheless, the social interaction behavior is the most diversified and, usually, most complex (in terms of the analytical tools to be used) category of crowd behaviors.

C. Leader–Follower Behavior: In certain scenarios, the behavior of a specific individual within a group of people is reflected in the uniform behavior of the whole group. Such motion patterns are considered leader–follower behaviors. In each group within a crowd, a person who initiates a pattern of motion followed by the other members of that group is identified as a leader of that group, and the other members as followers. In some applications of socio-cognitive crowd behavior analysis, in order to understand the behavior of each group in the crowd or the whole crowd as one group in the scene, it is critical to identify and localize the individual with the influential and impactful behavior on the crowd at that period of time.

Therefore, leader–follower can be considered a specific case of the social interaction behavior. Nevertheless, we believe it should be handled separately. First, this is a behavior of a particular importance in some common applications and scenarios such as evacuation, forecasting and guidance planning. Based on our study, there is restricted literature available in the public domain, and given the focus of these works being simulation-based analysis, we believe this category will trend more in the future of automated visual surveillance. Secondly, the algorithmic principles [12] of this behavior detection are (unlike in social interaction behavior in general) rather straightforward. In the leader–follower case, only one individual has an influence on a group, compared to the social interaction case where all members of a group are influential. Having such an assumption particularly simplifies the task of analyzing the scene in applications such as the aforementioned and limits the number of possible behaviors to be localized.

## II. LITERATURE REVIEW

### 2.1 Crowd Detection Steps

The detection of unexpected behaviors is difficult in crowds due to occlusions, and to be efficient it depends also on crowd density and available computing time. There are two different approaches, the object-based approaches which take each person as an object in the crowd, and holistic approaches which take the crowd as a single object. Holistic methods usually still have some occlusion problems and are sensitive to perspective variations. The main methodologies used for crowd behavior analysis are histograms of motion directions, optical flow, crowd density estimation and person tracking. These approaches will be discussed in more details bellow.

A) Pre-processing and selection of areas of interest: The analysis of the behavior of people and groups relies mainly on the analysis of motion in the video. Before proceeding to motion based solutions, it is common practice to define or estimate first regions where motion is supposed to occur. These areas are also called regions of interest. A region of interest may be a point, a line, a random shape, or an image set. It can be manually defined or estimated automatically using for example back ground extraction methods or by detecting points of interest. The simplest solution to delimit the area of interest is a manual definition of a movement zone. This is a practical solution in situations where the movement zones are known a priori and do not cover the entire scene. Also a motion area can be defined as a line (in the case of flow estimation) or as a mask defining motion zones. A motion mask can be used in conjunction with other motion zone detection methods to restrict further processing areas to specific locations.

Background modeling is used in different applications such as video surveillance or multimedia. The easiest way to separate the front of the background is to acquire an image of the background that contains no moving objects. A subtraction of images using an estimated threshold is then carried out between each new image acquired and the initial background image. However, a single background image is often not enough or unavailable because it is constantly modified by various events (e.g. lighting changes, objects entering or exiting the scene, etc.). In order to overcome the problems of robustness and adaptation, numerous works on background subtraction have been proposed based on Gaussian Mixture Models (GMMs), CNNS [13], etc. In order to improve further the preprocessing processes, points, objects or areas of interest (e.g. cars, lines, etc.) that represent and provide useful information may be considered too. In general, points of interest are pixels in an image that are likely to be tracked effectively in the video sequence (e.g. Harris corner detector).

B) Motion-based crowd behavior analysis: Optical flow approaches are based on the first degree Taylor series expansion of the change at each pixel. Given two images *I(x,* y, tn) and I(x, y, tn+1), the optical flow is defined under the assumption of brightness constancy as:

$$I(x, y, t) = I(x + dx, y + dy, t + dt).$$

Where $dx$ and $dy$ correspond to the motion vector over $dt$. Considering now the Taylor series expansion we obtain

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \partial I/\partial x \ dx + \partial I/\partial y \ dy + \partial I/\partial t \ dt + \ldots$$

Optical flow[14] is used extensively in crowd behavior analysis and the proposed solutions are considered holistic approaches. Therefore, it is not expected that one may to detect or segment each person of the crowd but the crowd is considered as a single entity. The advantage of holistic approaches based on optical flow is that computational complexity remains relatively low. Andrade et al. presented a framework for crowd behavior analysis that is based on optical flow information from video sequences aiming to represent the crowd behavior as flow variations over time. The flow features are encoded with Hidden Markov Models to allow for the detection of emergency or abnormal events in the crowd. In order to increase the detection sensitivity, a local modeling approach is used. The results with simulated crowds show the effectiveness on detecting abnormalities in dense crowds. Similar approaches considering flow information and holistic representations were proposed.

Figure 2.1: Left) Image from the original crowd. Right) Color coded representation of the optical flow.

C) Crowd Behavior and Density Analysis

In crowd analysis people counting and density estimation are important problems and since crowd density is one of the basic descriptions of the crowd status, it can provide significant information about behaviors and certain events. It would be a great help to make appropriate decisions for emergency and safety controls. Furthermore, it can be used for market analysis and, specifically, it could be used for developing service providers in public places, or supplying the current state and queues of waiting customers. There are two main approaches to calculate the crowd density. The direct approach is to segment and detect each person in the crowd, but this method is sensitive to occlusions. In the second approach people counting is carried out normally using the measurements of features with learning mechanisms or based on a statistical analysis of the whole crowd, aiming to achieve accurate counting estimates.

D) Person detection and tracking in crowded scenes

Another way to analyze crowd behavior is to track every person in the crowd. These methods have two main issues to overcome. First, the computation time increases significantly with the number of people that it is expected to track. Secondly, tracking is very sensitive to occlusions and therefore it operates well only for a low density crowd.

Figure 2.2: Bottom-up approaches for human detection and pose estimation in crowded scenes.

Crowd analysis becomes very popular research topic in the area of computer vision. A growing requirement for smarter video surveillance of private and public space using intelligent vision systems which can differentiate what is semantically important in the direction of the human observer as normal behaviors and abnormal behaviors. People counting, people tracking and crowd behavior analysis are different stages for computer based crowd analysis algorithm. In existing method following problems were encountered:

1) In existing method the optical flow for each frame is calculated using Lucas-Kanade algorithm [15]. This algorithm is well suited for local optical flow.

2) The Local optical flow suffers from Pixel correspondence problem. Given a pixel in the first image, look for a "nearby" pixel in the second image with the same brightness. (Brightness constancy).

3) This approach faces considerable complexity in detection of objects, tracking trajectories, and recognizing activities in dense crowds where the whole process is affected by occlusions.

Aiming at the above problems, this paper puts forward the characteristics of crowd movement based on optical flow field and Harris corner. Three different characteristics were extracted: the average kinetic energy of the population, the movement direction entropy and the population distance potential energy. In addition, two classifications of extreme learning machine and support vector machine are used to analyze the abnormal behavior of the crowd.

The overall algorithm structure of this dissertation can be divided into three steps:

1) Using the group movement feature algorithm to extract the average kinetic energy of the crowd, the movement direction entropy and the crowd distance potential energy can be used.

2) The extracted group movement feature are used as training samples, set up the data set and establish the abnormal event detection model through proposed algorithm.

## III. PROPOSED METHOD

Normally, the direction and speed of crowds are similar. However, when an abnormal event occurs, people will run away quickly due to fear to avoid potential danger. Abnormal behavior of the crowd has the characteristics of fast movement speed, sudden increase in acceleration, obvious concentration of movement in a certain direction or balance in multiple directions, and chaotic trajectory. As we all know, once the position of the pedestrian in the crowd is obtained, long-term tracking can be performed and the trajectory of the target can be calculated. However, due to the high density of pedestrians in the crowd and severe occlusion, it is impossible to track directly and accurately. Therefore, deep optical flow is introduced into counting the overall motion trajectory. Recently, the optical-flow calculation method based on convolutional neural network is the state-of-the-art algorithm. FlowNet uses a supervised method for deep learning training on the optical-flow prediction problem, and it is the first successful attempt to directly predict optical flow using a convolutional neural network. As it involves pixel changes and predictions, the input of the optical-flow network is usually a pair of images, and the output is the corresponding optical-flow graph. Flow-Field refers to the two-dimensional field obtained by projecting the instantaneous speed of the corresponding pixel connection in the image-pair. The movement direction of the pixel connection is naturally distributed in the horizontal and vertical directions on the plane. The optical-flow map is an image representation of the optical-flow field. Obviously, the optical-flow map is a dual-channel image. The final output performance is that different colors mean the direction of movement, and the color depth means the speed of pixel movement.

In our proposed work, we particularly focus on global abnormal crowd behaviour detection. In general, there are two main approaches for understanding global crowd behaviors: object-based and holistic approaches. In object-based methods, a crowd is considered as a collection of individuals. It is necessary to detect and track each individual to understand the crowd behavior. On the other hand, holistic approaches consider the crowd as a global entity and extract features such as using optical flow to represent the state of motion in the whole frame for higher level analysis.

In our research, we particularly focus on global abnormal crowd behavior detection in surveillance videos such as people suddenly start to run around in the same or different directions. Anomaly detection, also named as outlier detection, refers to detecting patterns in a given dataset that do not conform to an established normal behavior.

The framework of different phases is shown in figure given below:

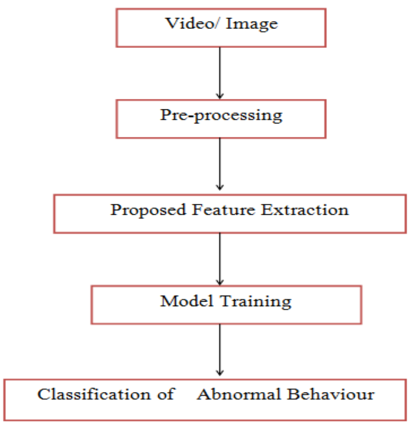
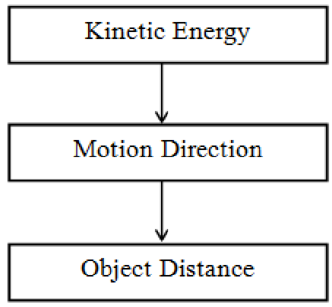



Figure 3.1: Proposed Framework.

The very first step is video capturing that is usually obtained from standard video sources such as online CCTV or off-line video databases. Before extracting features such as calculating the optical flow, some pre-processing operations such as background subtraction can be implemented, in order to get smoother feature and save computational time. Besides normal features, **proposed system** will extract some other relevant features also. Once features are extracted, the feature values sometimes need to be modeled, because it is not sufficient to detect the abnormal behavior with low level raw features in most of the time. Next with the modeled feature data the crowd or individual behaviors can be analyzed by proposed classifier.

### 4.2 Proposed Feature Extraction

Proposed system will extract multiple object movement features from video about objects shown in figure below:




- The distance of moving objects in video is expressed by average of kinetic energy, and defines frame's average kinetic energy $E(x)$ of the group:

$$E(x) = \sum_{i=1}^{n} m_i \frac{v_i^2}{N}$$

Where N is the number of the current frame Harris point (the total number of motion vectors), mi is the weight coefficient, and *vi* represents the *i* th motion's optical flow vector magnitude.

- The directional uniformity of the moving object in video is expressed by the direction entropy.
- The distance information of the motion of the moving object in video is represented by the distance potential energy, and the distance potential energy is calculated by calculating the Euclidean distance.

**Proposed Algorithm:**

**Step 1:** Take a video as Input.

**Step2:** Convert Video to frames.

**Step3:**Pre-processing the frames. Pre-preparing is an improvement of the picture information that stifles undesirable contortions or upgrades some picture highlights critical for further preparing.

**Step 4**: Histogram equalization. Here the basic idea is to change the original image pixel gray value of the number of pixels in the image gray value to widen, while the number of pixels in a small gray level reduction, the image is converted into the form of a histogram.

**Step5:** Image segmentation. It is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super-pixels). Here we are using edge detection technique for finding the boundaries of objects within images and in this research work we have used canny edge detector.

**Step6:** Morphological: This strategy confirms the picture with a little layout called organizing component. This organizing component is connected to every single imaginable area of the information picture and creates a similar size yield. In this method the yield picture pixel esteems depend on comparative pixels of info picture with is neighbors. This activity delivers another double picture in which if test is fruitful it will have non-zero-pixel esteem at that area in the information picture. There is different organizing component like jewel molded, square formed, cross molded and so on.

**Step7:** Object recognition: It is the way toward discovering examples of certifiable items, for example, human, bikes, and structures in pictures or recordings. Item location calculations commonly use extricated highlights and learning calculations to perceive occasions of an article like human movement examination. It is generally utilized in human movement in group region.

**Step8:** Classification: It has been applied to pixels or to images. When classifying pixels, we try to decide whether a given pixel belongs to a class. Detection is all about searching for objects or features within images and determining their locations. For example, finding the faces, signs or license plates in an image. When classifying images, we are trying to identify the type of image we have, such as what the main subject of the picture is (person, dog, bus, and building). We have used support vector machine. It is a powerful tool for binary classification, capable of generating very fast classifier functions following a training period. SVM can be considered as a linear method in a high-dimensional feature space nonlinearly related to the input space. Using kernels, all input data are mapped nonlinearly into a high-dimensional features space. It is possible to compute the separating hyper planes without explicitly carrying out the map into the feature space and classify to abnormal and normal images.

## IV. RESULTS

In this paper we are predicting that a Video is "Normal" or "Abnormal". We have a dataset that contains videos recorded by cameras and CCTV. Here we are predicting that if a image frame is not in comparison, then it will be classified as "Abnormal" and will predict as abnormal behavior.

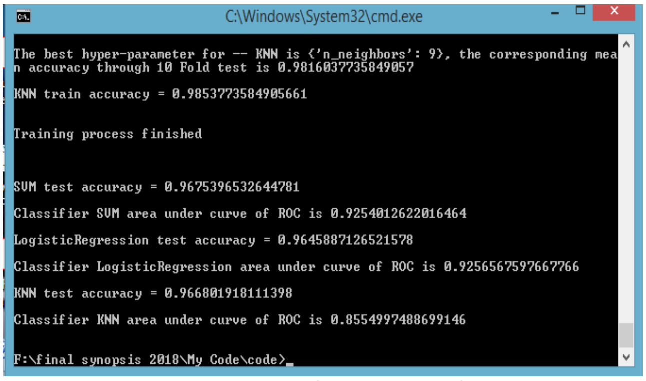


Figure 4.1: Snapshot of proposed classification.


Figure 4.2: Snapshot of proposed frames converted from video.

Table 1: Performance Evaluation.

| Method | Testing Accuracy (%) |
|---|---|
| **Logistic Regression** | **96.45** |
| **KNN** | **96.68** |
| **Proposed Method** | **96.75** |

It is observed that proposed classifier achieved best accuracy for classifying abnormal behavior of human in a crowd.

## CONCLUSION

We have presented an overview of recent and state-of-the-art techniques for crowd behavior analysis considering both fixed and moving cameras. The main research topics that were presented are crowd behavior analysis based on motion information, density estimation considering person detection and tracking methodologies, and low level pattern based solutions. Due to its importance in applications related to security and market analysis, anomaly detection in crowded scenes was discussed and state of the art approaches were presented.

Furthermore, advances in ML-based methods for crowd counting and density estimation are reviewed focusing on density estimation and the detection of abnormal events. Finally, evaluation metrics for crowd behavior analysis are presented and datasets captured from fixed and moving cameras for crowd analysis are summarized.

## *References*

[1] Luber, M., Stork, J.A., Tipaldi, G.D., Arras, K.O., 2010. People tracking with human motion predictions from social forces, In: Robotics and Automation (ICRA), 2010 IEEE International Conference on, pp. 464–469.

[2] Zhao, J., Xu, Y., Yang, X., Yan, Q., 2011. Crowd instability analysis using velocity-field based social force model. In: Visual Communications and Image Processing, VCIP. IEEE, pp. 1–4.

[3] Mazzon, R., Poiesi, F., Cavallaro, A., 2013. Detection and tracking of groups in crowd, In: Advanced Video and Signal Based Surveillance, AVSS, 10th IEEE International Conference on, 2013, pp. 202–207.

[4] Durupinar, F., Gudukbay, U., Aman, A., Badler, N.I., 2016. Psychological parameters for crowd simulation: From audiences to mobs. IEEE Trans. Vis. Comput. Graphics 22 (9), 2145–2159.

[5] Baig, M.W., Barakova, E.I., Marcenaro, L., Regazzoni, C.S., Rauterberg, M., 2014. Bio-inspired probabilistic model for crowd emotion detection, in: 2014 International Joint Conference on Neural Networks, IJCNN, pp. 3966–3973.

[6] Ali, S., Nishino, K., Manocha, D., Shah, M., 2013. Modeling, simulation and visual analysis of crowds: A multidisciplinary perspective. In: Modeling, simulation and visual analysis of crowds. Springer, pp. 1–19.

[7] Wang, T., Wang, J., 2017. Optimization of evacuation path decision with bim and psychological support, in: 2017 29th Chinese Control and Decision Conference, CCDC, 814–821.

[8] Liu, W., Lau, R.W., Manocha, D., 2016. Robust individual and holistic features for crowd scene classification. Pattern Recognit. 58, 110–120, (Suppl. C).

[9] Fradi, H., Luvison, B., Pham, Q.C., 2017. Crowd behavior analysis using local mid-level visual descriptors. IEEE Trans. Circuits Syst. Video Technol. 27 (3), 589–602.

[10] Y. Cong, J. Yuan and J. Liu, "Abnormal Event Detection in Crowded Scenes using Sparse Representation," Pattern Recognition, vol. 46, no. 7, pp. 1851-1864, 2013.

[11] B. Solmaz, B. E. Moore, and M. Shah, "Identifying behaviors in crowd scenes using stability analysis for dynamical systems", IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 10, pp. 2064–2070, Oct. 2012.

[12] Bouwmans, T., Silva, C., Marghes, C., Zitouni, M.S., Bhaskar, H., Frelicot, C., 2018. On the role and the importance of features for background modeling and foreground detection. Comput. Sci. Rev. 28, 26–91.

[13] J. R. Model and A. Savakis, "Anomaly detection in video using predictive convolutional long short-term memory networks," arXiv preprint arXiv: 1612.00390, 2016.

[14] Hsin Chun Tsai, "The optical flow-based analysis of human behavior specific system". Orange Technologies (ICOT), 2013 International Conference on Year: 2013.

[15] Bera, S. Kim and D. Manocha. Real time anomaly detection using trajectory-wise crowd behavior learning. In Proceedings of International Conference on Computer Vision and Pattern Recognition, CVPRW'13, pages 50-57. IEEE, 2016.