ABNORMAL HUMAN ACTIVITY DETECTION SYSTEM

Enrolment No(s). – 17803005, 17803011, 17803016

Name of Student – Priyanka Parashar,Aradhya Mathur,Mitushi Agarwal

Name of Supervisor – Dr. Adwitiya Sinha

December – 2020

For the partial fulfilment of the Degree of

5 Year Dual Degree Program in Btech

in

Computer Science

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING & INFORMATION TECHNOLOGY

JAYPEE INSTITUTE OF INFORMATION TECHNOLOGY, NOIDA

TABLE OF CONTENTS

DECLARATION

We hereby declare that this submission is our own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Place: Jaypee Institute of Information Technology

Date: 07/12/2020

Name: Priyanka Parashar, Aradhya Mathur, Mitushi Agarwal

Enrollment No - 17803005, 17803011, 17803016

Signature:

CERTIFICATE

This is to certify that the work titled "**ABNORMAL HUMAN ACTIVITY DETECTION SYSYTEM**" submitted by Priyanka Parashar, Aradhya Mathur, Mitushi Agarwal in partial fulfillment for the award of degree of 5 Year Dual Degree of Jaypee Institute of Information Technology, Noida has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Signature of Supervisor ………………………………………….

Name of Supervisor- Dr. Adwitiya Sinha

Designation-

Date -07/12/2020

ACKNOWLEDGEMENT

We wish to express our sincere gratitude *to Dr. Adwitiya Sinha* for mentoring our final year major project. We are grateful for his valuable guidance and encouragement in carrying out this project efficiently within the given time constraints.

Signature of the Student ……………………..

Name of Student Priyanka Parashar, Aradhya Mathur, Mitushi Agarwal

Enrollment Number 17803005, 17803011, 17803016

Date 07/12/2020

SUMMARY

In this project, we propose a novel method for unusual human activity detection in crowded scenes. Specifically, rather than detecting or segmenting humans, we devised an efficient method, called a motion influence map, for representing human activities.

The key feature of the proposed motion influence map is that it effectively reflects the motion characteristics of the movement speed, movement direction, and size of the objects or subjects and their interactions within a frame sequence.

Using the proposed motion influence map, we further developed a general framework in which we can detect both global and local unusual activities. Furthermore, thanks to the representational power of the proposed motion influence map, we can localize unusual activities in a simple manner.

___________________ ___________________

Signature of Student	Signature of Supervisor
Name:Priyanka Parashar, Aradhya Mathur, Mitushi Agarwal	Name:Dr.Adwitiya Sinha
Date 07/12/2020	Date 07/12/2020

LIST OF FIGURES

1. Introduction

1.1 General Introduction

With the expansion in crime and abnormal human activity that has been occurring, security has been given the most extreme significance recently. Numerous associations have introduced CCTVs for the consistent checking of individuals and their collaborations. A large portion of the memory spaces of the business is involved in big data. The execution of CCTV cameras in all areas because of security purposes and utilization of CCTV cameras is basic however, it devours more memory spaces to store information. Security is utilized for robbery recognizable proof, viciousness recognition, unapproved people entering, criminal behaviour in a locale. Thus, for all unusual movement's security assumes a significant job, so security must be actualized in the area of more privacy. Utilizing CCTV videos in the past day's strategy to discover the robbery happenings and different exercises, this is a dreary cycle and tedious job. The crowd may emerge in occupied roads, games, music shows, and fights, among others. Although individuals in a crowd regularly move in an organized way, little unsettling influences may prompt a frenzy circumstance and perhaps grievous results.

Abnormal Human Activity Detection System is a model for unusual human activity identification in a crowded scene. The development heading of a walker inside a group can be affected by different factors, for example, impediment along the way, close by walkers, and moving trucks. This communication interaction, which we call the "motion influence". In particular, rather than distinguishing or dividing people, we integrate a productive strategy, called a motion influence map, for characterizing to human movement. The key element of the proposed motion influence map is that it viably mirrors the movement qualities of the motion speed, motion heading, and size of the articles or subjects and their interaction inside a frame. Utilizing the proposed motion influence map, we further built up a general system in which we can distinguish both global and local uncommon exercises. Anomaly recognition and limitation can be separated into two sub-issues: 1) how to portray crowd practices, and 2) how to quantify the "anomaly score" of particular conduct. For the main issue, we propose to show movement designs in crowd through the utilization of a mixture of dynamic textures (MDT), which is a bound together portrayal catching both the appearance and dynamics of visual cycles. In the subsequent part, rather than legitimately demonstrating the atypical conduct itself, the routineness is first learned, and afterward the "anomaly score" of perception is processed by estimating the distinction from the regularity model. In particular, two segments are proposed to mirror the regularity in alternate points of view. Moreover, with a motion influence map, we can basically localize abnormal exercises. What makes Abnormal Human Activity Detection System model different is that Existing methodologies center particularly around motion information, overlooking anomaly information because of varieties of item appearance. This makes them impenetrable to irregularities that don't include motion anomalies. Besides, descriptors, for example, optical flow, pixel change histograms, or other conventional foundation deduction activities, are hard for swarmed scenes, where the foundation is by definition dynamic, of inescapable mess, and confounded impediments.

1.2 Problem Statement

With the increase in the amount of anti-social activities taking place in the environment, security has been given the utmost importance lately. Therefore, organizations require a constant monitoring of people and their interactions. Since this constant monitoring of data by humans to judge if the events are abnormal is a near impossible task as it requires a lot of workforce and constant attention. Therefore, the challenge that comes up is the demand for an automatic and intelligent analysis for such video sequences. Our project comes forward as an attempt to provide solution to such a problem as the model developed is a smart surveillance system which can detect unusual or abnormal activity automatically. A method for representing the motion characteristics is described for detection and localization of unusual activities in the crowd scenes on a generalized framework which includes both a local and global range for detection of such activities.

1.3 Novelty of the Problem

We have chosen a project based on a significant usage in day to day fields as with the increase in the number of anti-social activities that have been taking place, security has been given utmost importance lately. Many organizations have installed CCTVs for constant monitoring of people and their interactions. Since constant monitoring of data by humans to judge if the events are abnormal is a near impossible task as it requires a workforce and their constant attention. This creates a need to automate the same. Also, there is a need to show in which frame and which parts of it contain the unusual activity which aid the faster judgment of that unusual activity being abnormal. Therefore, with the help of the created model we tried to put forward an attempt to provide solution to such a problem as the model developed is a smart surveillance system which can detect unusual or abnormal activity automatically.

1.4 Brief Description of Solution Approach

To develop an accurate human activity analysis system that would assure maximum accuracy in detection of abnormal or unusual activities of people in a crowd automatically. This system can be effectively installed in a CCTV Camera to perform vison-based surveillance on a video footage. A method for representing the motion characteristics is described for detection and localisation of unusual activities in the crowd scenes on a generalised framework which includes both a local and global range for detection of such activities.

2. Literature Survey

2.1 Summary of papers studied

There are various existing surveys that help in providing a better understanding of detection of unusual activities, vision-based surveillance, intelligent video surveillance systems, crowded scenes, motion-influence map that helps us in detecting abnormal activities in general as well as crowded scenes. Reuben A. Farrugia in [1] explained how there are numerous amounts of CCTV cameras installed which requires a lot of human labor to monitor them and sensor's output is too expensive. So, he presents an idea of vision-based surveillance from a video footage which will detect and tracks vehicles and humans for the video. Waleed Albattah in [2] explained how crowd analysis in recent years have been in active research areas. Over the past few years what all proposed methods have been used for crowd analysis and why there is need for it, and what are the applications of it. In his paper he presented a detailed review of state-of-art method used for crowd analysis and management and what the advantages and disadvantages of state-of-art method. Mariem Gnouma in [3] proposed two methods to detect and track unusual events in scenes. By using the Gaussian mixture model (GMM) technique for collecting statistical model of each individual for time tracking of people and improved this to IGMM to detect and track the individuals in crowd. Proposed a descriptor for detection of anomalous events in the video surveillance named the Distribution of Magnitude of Optical Flow (DMOF). This descriptor algorithm is based on adjustment velocity field by operating the intensity of light. Mikel D. Rodriguez in [4] described an approach for detecting and segmenting humans in a crowded video sequence on the basis of posture of human beings. They learned the postures and cluster them on the basis of various human postures. The first phase is detecting the human body in the video sequence and the second phase is segmenting on basis of body posture clusters. So, the proposed method is effective yet efficient for detecting and segmenting humans in crowded scenes where a diverse set of activities are being performed with lots of people. Waqas Sultani in [5] proposed a Multiple Instance learning (MIL) model which is used to detect video segments as anomaly or non-anomaly by giving anomaly scores. He trained the model on both anomaly videos and normal videos. The dataset he is using has 128 hours of videos which contains 1987 real world untrimmed surveillance video which has both anomalies videos such as fighting, car accident, robbing, etc and normal videos. This method of MIL shows betters results and improvements than state-of-art method. The recognition performance is low because of the challenging dataset which opens it up for more future work. Sunil Malviya in [6] says Dubious conduct is perilous out in the open regions that may cause weighty causalities. There are different frameworks created based on video outline securing where movement or walker identification happens yet those frameworks are not wise enough to recognize the surprising exercises even at continuous. Here framework utilizes the OpenCV library for characterizing diverse sorts of activities at ongoing. The motion influence map has been utilized to speak to the movement examination that oftentimes changes the situation starting with one spot then onto the next. The framework utilizes a pixel-level introduction for making it simple to comprehend or distinguish the genuine circumstance. Alisha Ahir in [7] summarizes strategies that are accessible in the field of the discovery of strange exercises and studies about the different techniques, methods and examine their points of interest and hindrances. Furthermore, to analyse these advances dependent on their presentation, there are numerous strategies like Spatio-temporal Saliency Detection, Graph Formulation, Arranged GMM, based upon Sparse remaking, and utilizing Measurable Hypothesis Detector and so on. Dayana R [8] in industry enormous information applications are devouring the majority of the spaces. Among certain instances of large information, in light of the security reason CCTV cameras are actualized in all spots where security has a lot of significance. Security can be characterized in various manners like burglary distinguishing proof, viciousness discovery, and so on, and talks about recognizing and perceiving the facial highlights of the people utilizing deep learning ideas. This paper incorporates deep learning in ideas that begins from object location, activity recognition, and distinguishing proof. The issues perceived in existing techniques are recognized and summed up. M.H. Sedky in [9] There are quick business requirements for smart video observation frameworks that can utilize the current camera organization (for example CCTV) for more quick security frameworks and to contribute in more applications (alongside or) as opposed to security applications, This work presents another grouping for smart video inspection frameworks contingent upon their business applications and features various connections between the examination and the business applications. The work detailed has both exploration and business inspirations. Our objectives are first to characterize a conventional model of smart video observation frameworks that can meet prerequisites of solid business applications. Our subsequent objective is to sort extraordinary smart video observation applications and to relate the abilities of computer vision calculations to the necessity of business application. Igor R. de Almeida [10] Groups emerge in an assortment of circumstances, for example, public shows and matches. In ordinary conditions, the mass moves in an efficient way, however, alarm circumstances may lead to calamitous outcomes. We propose a computer vision technique to recognize movement design changes in human groups which can be identified with an uncommon occasion. The proposed approach can distinguish worldwide changes, by assessing 2D movement histograms in time, and neighbourhood impacts, by recognizing bunches that present comparable spatial areas and speed vectors. Kosuke Hara in [11] has tried to describe a model that can be used for detecting human behaviour in an intelligent house by the means of sensors. He has formulated a Markov Chain Model to show the validity of the state transition probability and duration time distribution through the detection of unusual activity or unusual behaviour in three different sets of human behaviour data. The results indicate that a probabilistic model can be effectively used to describe the daily human activity and a successful detection of its unusual behaviour can also be done. Zafar A. Khan in [12] has summarized a Human Activity Recognition System that can we developed to improve the lifestyle and conditions of the elderly people by observing their daily life so that they can survive independently in the world, The dataset used for detecting and validating the same includes mainly six abnormal activities for instance the forward fall, backward fall, chest pain etc. The major algorithms that highlights the model are R-transform and KDA which are used the curb the problem of the distance that changes continuously with the movement of the person and differentiating high similarities in various abnormal activities respectively. The model turns out to have an average accuracy of 95.8% and is said to be effectively used for the health care of elderly people at home. Liang Chen in [13] has effectively tried to develop a Human Pose recognition system which is counted as an important study in the computer vision and pattern recognition field. The major task of reduction of misjudgements and confusion that occur due to similar movements and external factors during gesture recognition has been effectively creating a DBLSTM model which is developed using deep learning and data fusion. The average recognition rate of the model turns out to be 92.7% and hence it can be used to solve the confusion problem for similar human activities. Zhuang Miao in [14] focuses on one of the important research topics in the computer vision that is the moving object detection and tracking technology. This filed is widely applied for public safety inspection, traffic monitory etc. An idea of developing an intelligent video surveillance for the same has been proposed, basically for analysing and detecting suspicious event that includes all kinds of security threats. The paper is concluded by presenting an automatic alarm sample for video surveillance. Edi Noersasongko in [15] provides us with the comprehensive and systematic review of the current scenario of the video surveillance system. It gives us information about the current ongoing state of this technology along with the research that has already been done so far in the same field which includes the attempt to integrate the computer vision and other related intelligence applications of video surveillance together. Krishna Reddy Konda in [16] has tried to describe an algorithm for moving object detection and segmentation, operating on H.264-bit streams. Compared to more traditional pixel-based approaches, the novelty of the algorithm consists of directly using the motion features embedded into the H.264-bit stream, thereby achieving real time operational capability. This makes the algorithm ready to be installed in any video surveillance system, enabling for better resource allocation and facilitating the deployment of distributed systems. The method we propose measures the statistical disorder of the motion field at the boundary of the moving objects, achieving at the same time detection and segmentation. In order to refine the segmentation, results, the temporal correlation of motion vectors is analysed. The algorithm has been tested on the traditional videos used to benchmark video compression algorithms, as well as on a subset of sequences from the I-lids dataset, to demonstrate its generalization capabilities. Henan Guo in [17] has summarized that how with the rapid development of science and technology and with the continuous improvement of people's safe sense, Video Surveillance System has already been widely used in several fields, such as military affairs, production, daily life and so on. As the most basic and important part of Video Surveillance System, the detection of moving objects has been paid more attention. Therefore, this paper mainly implements and analyses several commonly used methods for moving object detection in Video Surveillance System, such as temporal difference, median filtering and Single Gaussian model. The advantages and disadvantages of the methods are pointed out to help people make choices in applications. Abhishek Mohite in [18] has effectively tried to develop a novel method for unusual human activity detection in crowded scenes. Specifically, rather than detecting or segmenting humans, they have devised an efficient method, called a motion influence map, for representing human activities. The key feature of the proposed motion influence map is that it effectively reflects the motion characteristics of the movement speed, movement direction, and size of the objects or subjects and their interactions within a frame sequence. Using the proposed motion influence map, they further developed a general framework in which we can detect both global and local unusual activities. Furthermore, thanks to the representational power of the proposed motion influence map, it can localize unusual activities in a simple manner. They performed experiments on three public datasets, and compared the performances of the proposed method with that of other state-of-the-art methods, and showed that the proposed method outperforms these competing methods. Video surveillance is a prominent area of research which includes recognition of human activities and categorisation of them into usual (normal), unusual (abnormal) or suspicious activities. Due to exponential increase in crime rate, surveillance systems are being put up in malls, stations, schools, airports etc. S. R. Balaji in [19] focuses on one of the important research topics in the computer vision that is Image processing. It is a method of extracting some useful information by converting image into digital inform by performing some operations on it. Object detection and tracking are the task that is important and challenging such as video surveillance and vehicle navigation. Video surveillance is a technology which works in dynamic environment in various events such as sports, public safety, and management of traffic. This paper reviews the various challenges and aspects of detection and tracking of objects. Kyungbok Min in [20] provides us with the comprehensive and systematic review of the current scenario of the Human activity recognition (HAR) technology that analyses data acquired from various types of sensing devices, including vision sensors and embedded sensors, has motivated the development of various context-aware applications in emerging domains, e.g., the Internet of Things (IoT) and healthcare. Even though a considerable number of HAR surveys and review articles have been conducted previously, the major/overall HAR subject has been ignored, and these studies only focus on particular HAR topics. Therefore, a comprehensive review paper that covers major subjects in HAR is imperative. This survey analyses the latest state-of-the-art research in HAR in recent years, introduces a classification of HAR methodologies, and shows advantages and weaknesses for methods in each category. Specifically, HAR methods are classified into two main groups, which are sensor-based HAR and vision-based HAR, based on the generated data type. After that, each group is divided into subgroups that perform different procedures, including the data collection, pre-processing methods, feature engineering, and the training process. Moreover, an extensive review regarding the utilization of deep learning in HAR is also conducted. Finally, this paper discusses various challenges in the current HAR topic and offers suggestions for future research. Hyun-Sang Park in [21] says Crowd anomaly detection is a key research area in vision-based surveillance. Most of the crowd anomaly detection algorithms are either too slow, bulky or power-hungry to be applicable for battery-powered surveillance cameras. In this paper, we present a new crowd anomaly detection algorithm. The proposed algorithm creates a feature for every super pixel that includes the contribution from the neighbouring super pixels only if their direction of motion conforms with the dominant direction of motion in the region. We also propose using univariate Gaussian discriminant analysis with the K-means algorithm for classification. Our method provides superior accuracy over numerous deep learning-based and handcrafted feature-based approaches. We also present a low-power FPGA implementation of the proposed method. The algorithm is developed such that features are extracted over non-overlapping pixels. This allows gating inputs to numerous modules resulting in higher power efficiency. The speed, power and accuracy performance of our method make it competitive for surveillance applications, especially battery-powered surveillance cameras. C.M. Patil in [22] says Human tracking & action recognition is an important research area in computer vision. Despite a lot of progress in the field, visual tracking remains a difficult problem due to many challenges. In this survey paper, a brief review of the research work that is being be done in the field of multiple human tracking & action recognition is being presented along with some of their drawbacks. V.M. Fanase in [23] says Problem of detection and localization of abnormal behaviours in crowded scenes is considered in this paper. To extract various activities of crowd, spatiotemporal Laplacian eigenmap method is proposed. Here, spatiotemporal variations of local motions are learned in an embedded space. Regular behaviour of a crowd is modelled using representatives of various activities. To detect abnormal behaviour in local as well as global contexts and to localize regions showing abnormality, model of regular crowd behaviour is used. Computational simplicity is the feature of this method. Pooja Shah in [24] says Crowd investigation turns into the most dynamic research in PC vision, in the shrewd video reconnaissance zone. In this research, an online solution is implemented to detect crowd abnormalities, which can be caused by panic. After calculating optical flow and calculating magnitudes using the Farneback method, an activity map is created using multiple frames to see the flow persistency overtime. The activity map is used to create two metrics, TOV (Temporal Occupancy Variation) and entropy, respectively. In this research, the measurements alongside an edge to the group are utilized. Subtleties will be examined in the accompanying segments. The movement directions are centred to watch the group and optical stream strategies are utilized to gain the streak lines and faculty directions. Also, the exhibition measurements are computed like accuracy FTR, TPR, and ROC. Ming Dong in [25] says Automatic detection of abnormal crowd activities is one of central tasks in video surveillance. In this paper we present a matrix approximation-based method to detect abnormal crowd behavior. In our approach, we model typical motions associated with normal crowd behaviours with a set of motion subspaces, computed through low-rank matrix approximation. Then, abnormal crowd behaviours are identified by the motion deviations from the representative subspaces. Our method does not require complicated tracking or classification method, and can fast detect abnormal events in complex crowd scenes. In addition, through the adaptive learning module, our model is built on the observed data, and can be expanded by incorporating new crowd behaviour patterns during the detection process. The results on simulated crowd scenes show the effectiveness of our method. Marco Bertini in [26] explained an approach for anomaly detection and localization, in video surveillance applications, based on spatio-temporal features that capture scene dynamic statistics together with appearance. Real-time anomaly detection is performed with an unsupervised approach using a nonparametric modelling, evaluating directly multi-scale local descriptor statistics. A method to update scene statistics is also proposed, to deal with the scene changes that typically occur in a real-world setting. The proposed approach has been tested on publicly available datasets, to evaluate anomaly detection and localization, and outperforms other state-of-the-art real-time approaches. Arun Kumar Jhapate in [27] describes suspicious behaviour is dangerous in public areas that may cause heavy causalities. There are various systems developed on the basis of video frame acquisition where motion or pedestrian detection occur but those systems are not intelligent enough to identify the unusual activities even at real time. It is required to recognized scamper situation at real time from video surveillance for quick and immediate management before any casualties. Proposed system focuses on recognizing suspicious activities and target to achieve a technique which is able to detect suspicious activity automatically using computer vision. Here system uses OpenCV library for classifying different kind of actions at real time. The motion influence map has been used to represent the motion analysis that frequently changes the position from one place to another. System uses pixel level presentation for making it easy to understand or identify the actual situation. Bruno Jobar in [27] presents a new approach for animating 2D steady flow fields. It is based on an original data structure called the Motion Map. The Motion Map contains not only a dense representation of the flow field but also all the motion information required to animate the flow. An important feature of this method is that it allows, in a natural way, cyclical variable-speed animations. As far as efficiency is concerned, the advantage of this method is that computing the Motion Map does not take more time than computing a single still image of the flow and the Motion Map has to be computed only once. Another advantage is that the memory requirements for a cyclical animation of an arbitrary number of frames amounts to the memory cost of a single still image. Shihui Huang in [29] describes Human activity recognition in videos is important for content-based videos indexing, intelligent monitoring, human-machine interaction, and virtual reality. This paper uses the low-level feature-based framework for human activity recognition which includes feature extraction and descriptor computing, early multi-feature fusion, video representation, and classification. This paper improves the first two steps.

We propose a spatio-temporal bigraph-based multi-feature fusion algorithm to capture the useful visual information for recognition. Meanwhile, we introduce a compressed spatio-temporal video representation to bag of words representation. Our experiments on two popular datasets show efficient performance. Adnan Khalid in [30] explains an abnormal behaviour detection algorithm for surveillance is required to correctly identify the targets as being in a normal or chaotic movement with the help of a model. The uniqueness of this algorithm is the use of foreground detection with Gaussian mixture (FGMM) model before passing the video frames to optical flow model using Lucas-Kanade approach. Information of horizontal and vertical displacements and directions associated with each pixel for object of interest is extracted. These features are then fed to feed forward neural network for classification and simulation. The study is being conducted on the real time videos and some synthesized videos. Accuracy of method has been calculated by using the performance parameters for Neural Networks. In comparison of plain optical flow with this model, improved results have been obtained without noise. Classes are correctly identified with an overall performance equal to 3.4e-02 with & error percentage of 2.5.

2.1 Integrated Summary

With the developing number of CCTV cameras introduced in private and public zones, there has been an interest in the automated and smart study of video patterns utilizing computers abnormal occasion or action detection in a crowded the scene has of late been of the incredible fascination for the zone of vision-based surveillance. unalike past techniques, which have concentrated on either local or global abnormal action observation, from our research we have tried to propose a technique for constituting motion features inside a frame to distinguish and restrict abnormal human activity in a jam-packed scene. Due to the represented ability of the proposed motion influence map for both realities, we can arrange a frame as usual or unusual, and confine the regions of abnormal exercises inside a frame. For a genuine application, a brilliant smart surveillance framework requires to effectively to distinguish both local and global abnormal exercises inside a system. In our study on a public dataset, we approved the viability of the proposed technique, which outdo other contending methods. The principle focal point of this work is to distinguish abnormal activity inside a crowded scene, for which the cameras normally spread a wide region, bringing about little items being present in the scene without a huge point of view changes. Additionally, our evaluation was restricted to a fixed perspective, and there is a restriction in the relevance of the methodology for observation cameras with pan, zoom, or tilt functionality. Right now, the proposed strategy manages static cameras. Although, it very well may be effectively stretched out to PTZ cameras utilizing localization results. In such a manner, we accept that there are no enormous scaling changes or pans, tilt, or zooms in a scene, which might be an expected restriction of our strategy. Managing these functionalities is a region of imminent exploration, which will be led by further broadening the proposed strategy.

3. Requirement Analysis and Solution Approach

3.1 Overall description of the project

With the increase in the number of anti-social activities that have been taking place, security has been given us most importance lately. Many organizations have installed CCTVs for constant monitoring of people and their interactions. For a developed country with a population of 64 million, every person is captured by a camera ~ 30 times a day. A lot of video is generated and stored for certain time duration (India: 30 days). A 704x576 resolution image recorded at 25fps will generate roughly 20GB per day. Since constant monitoring of data by humans to judge if the events are abnormal is a near impossible task as it requires a workforce and their constant attention. This creates a need to automate the same. Also, there is a need to show in which frame and which parts of it contain the unusual activity which aid the faster judgment of that unusual activity being abnormal. The method involves generating motion influence map for frames to represent the interactions that are captured in a frame. The main characteristic feature of the proposed motion influence map is that it effectively depicts the motion characteristics of the movement speed, movement direction, and size of the objects and their interactions within a frame sequence. Its further extracts frames of high motion influence values and compares with the testing frames to automatically detect global and local unusual activities.

3.2 Requirement Analysis

3.2.1 Functional Requirements

The system should process images in the chosen image formats. (jpg, png, bmp, tif) The system shall detect abnormal frames and display the exact time from the start at which the abnormal event is detected. The length of the test video can vary and no particular limit is imposed. In the presence of no abnormality, user should be displayed with a message that the video is normal. After the whole video is processed the user should be given an option to see the portion of the video where abnormal event was detected. Only the frames to be suspected as abnormal should be played as a video. Option to see the abnormal portion of the video any number of times should be given to the user.

3.2.2 Non-Functional Requirements

- **Usability**

  The system is easy to train and test thus navigates in the most expected way with less delay. Since the algorithm is written in such a way that a lot of parallel computation can be performed, extensively high frame rates can be achieved by inculcating multi-threading. User will be allowed to add his frames easily and direct testing can be started on those frames. Thus, it is user friendly, reducing complexity on users and getting them a better result in a faster way.

- **Assumption**

  Test frames are either in .tif, .jpg, .png, .bmp formats. Each frame is taken to represent one second of the video. So, a 200 second video is assumed to generate 200 frames. The learning rate for training is chosen to be 0.0001 and the acceptance value of error in training is set to 0.04. In the testing phase threshold for reconstruction error is assumed to be 0.00000045 for optimal accuracy during testing.

- **Performance**

  Pre-captured video frames will be analysed fast as a significant amount of time is saved in not capturing the video. On the other hand, real-time capturing will extend the required time. Detection of abnormality in the frames is done at a rate of 150fps using parallel processing of data in the testing phase. An accuracy of 90% is achieved with the sparse combination method implemented.

- **Reliability**

  The software is tested with varies dataset mentioned in this report and the output is very close to the actual abnormal scenarios thus turning out to be reliable when all the assumptions mentions are considered.

3.2.3 Experimental Work

The dataset has major attributes as:

- Crowd Escape Panic, 11 Videos, 3 Scenes, Videos: a normal starting section and an abnormal ending section
- We parceled each edge into 8×8 nonoverlapping squares, and set the edge to the most extreme highlight an incentive in the movement impact guides of the preparation pictures. The basis for this methodology is that we expect the unusual exercises to bring about higher movement impact esteems than the typical exercises.

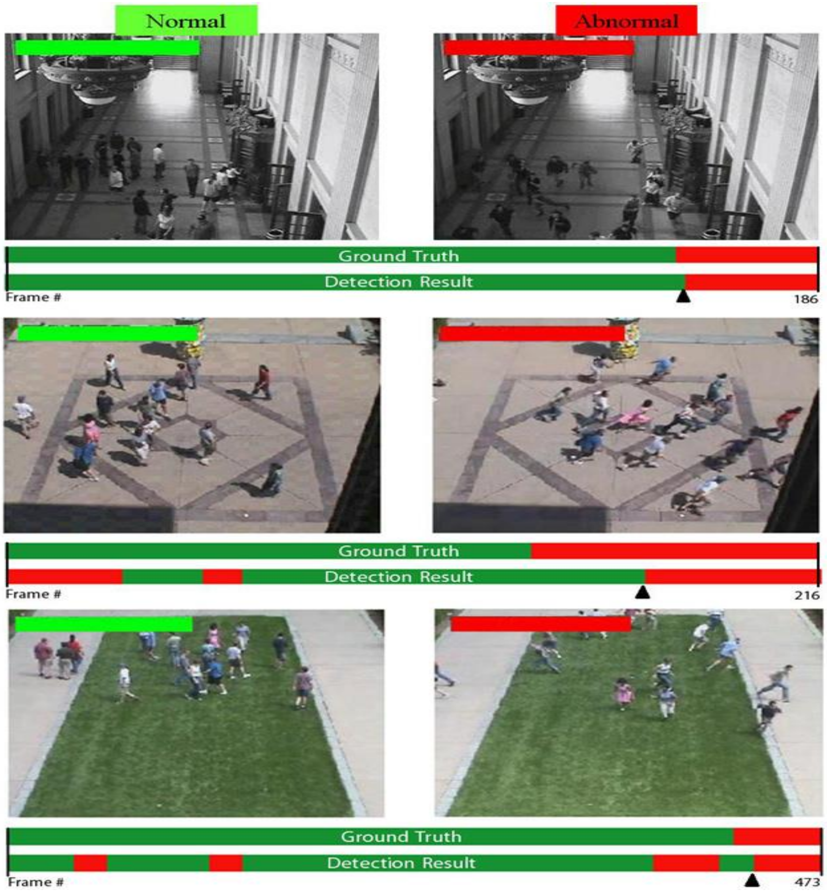


Figure 1:Normal vs abnormal frame

The second part of dataset:

- was obtained with a fixed camera mounted at a height, sitting above common walkways. The group thickness in the walkways was variable, running from inadequate to extremely swarmed. In the ordinary setting, the video contains just walkers.
- Anomalous occasions are expected to either:
- the flow of non-passerby substances in the walkways
- atypical walker movement designs

Peds1: clasps of gatherings of individuals strolling towards and away from the camera, and some measure of point of view mutilation. Contains 34 preparing video tests and 36 testing video tests.

(a)

(b)

Fig 2(a,b):Dataset scene 1

Peds2: scenes with passer by development corresponding to the camera plane. Contains 16 preparing video tests and 12 testing video tests.

(a)

(b)

Fig3(a,b):Dataset scene 2

3.3 Solution Approach

The Abnormal Human Activity Detection System is an intelligent open computer vision-based surveillance system. For the implement of this automated model effectively, we have surveyed various Research Papers in order to come forward with most accurate and efficient results.

3.3.1 Data Input and Pre-processing

For the pre-processing the input video file is subjected to the system. The video is split in series of images known as frames and these frames are handled in a sequence. The frame which is RGB is correspondingly converted to gray scale. The images which are gray scaled they contain only the intensity details instead of apparent colours. The vector which is gray scaled is one dimensional on the other hand vector RGB is three dimensional (it contains red, green and blue colours).

RGB scale(a)

Gray Scale image(b)

Fig4(a,b):RGB to gray scale image

3.3.2 Optical Flow Analysis

After the pre-processing step which was done for each frame in the given video, the next step is Optical Flow which we determine for every pixel of frame by the use of FreneBack algorithm. Optical flow is the example of clear movement of items, surfaces, and edges in a visual scene brought about by the general movement between a spectator and the scene. The optical flow vector has the form (r, θ), where the immensity of each pixel is represented by r, and the θ is representing the direction by the where every pixel has moved comparative with the relating pixel in the past frames. The dense optical flow in open cv is computed by the use of Gunnar Farneback's algorithm through the calcOpticalFlowFarneback( ) function.

After calculating the optical flow for each and every pixel inside the frame we divide the frame into blocks. Without the loss of generality, the frame is divided into M x N equal blocks. The optical flow for each block is calculated as the average of optical flow of all the pixels within a block. We do this after we have split up the frames into blocks. How much any frame has proceeded and in which direction when compared to the block of earlier frame is calculated using the vector (r, θ) of the optical flow of block.

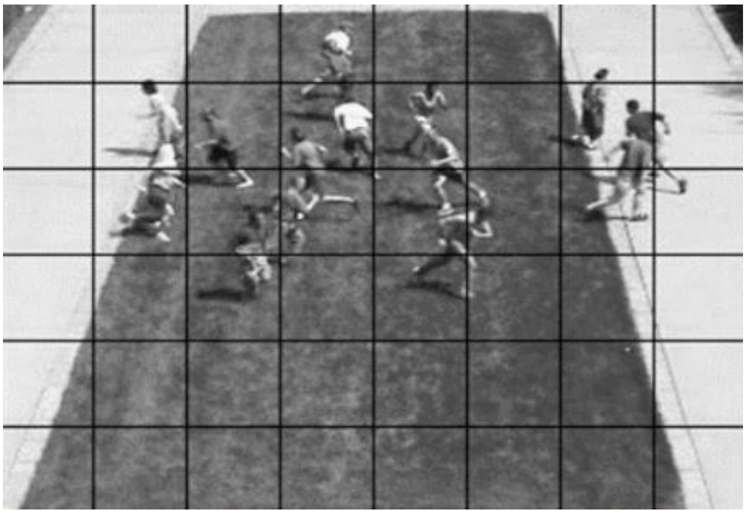


Fig 5 : Image divided into blocks.

3.3.3 Motion Information Using Motion Influence Map

The motion direction of a pedestrian inside a crowd can be affected by different variables, for example, obstructions along the way, close by walkers, and moving trucks. This interaction of attributes is known as motion influence. We accept that the squares under impact to which a moving item can influence are dictated by two components:

1. The direction of movement.
2. The speed of the motion. The quicker a body moves, the additionally neighbouring blocks that are affected by the blocks. Neighbouring blocks have a higher impact than far off blocks.

Feature Extraction:

In motion influence map the feature extraction, a block where an unordinary action happens, alongside its neighbouring blocks, has distinctive vectors of motion influence. Moreover, since a movement is caught by different back to back frames, we extricate a feature vector from a cuboid characterized by $n \times n$ blocks over the latest t number of frames.

Making Mega blocks frames are parcelled into non-covering mega blocks, every one of which is a blend of different motion influence blocks. The Motion Influence estimation of a Mega block is the summation of motion influence estimations of littler blocks comprising a bigger block. Taking out features after the ongoing 't' number of frames are isolated into Mega blocks, for every mega block, an $8 \times t$-dimensional connected component vector is extricated over all the frames.

Fig 6 : Motion Influence map

Clustering:

Clustering of every mega block, clustering is performed using the spatio-temporal highlights and set the focuses as codewords. In our training stage, we utilize just video clasps of ordinary activities. Accordingly, the codewords of a mega block model the examples of regular activities that can happen in the separate zone.

3.3.4 Testing Phase

After the generation of codewords for normal activities, it's time for the testing phase. Basically, the time has come for testing the model which we have created on the dataset which contains both normal and abnormal activities.

After the taking out the features vector of spatio-temporal for each and every mega block. A minimum distance matrix (E) is constructed above mega block in which the estimation of a component is characterized by the base Euclidean separation between an element vector of the current test outline and the codewords in the comparing mega block.

3.3.5 Frame and Pixel level detection for Abnormal Activities

Frame level observation of anomalies in a minimum distance matrix, the littler the estimation of an element, the less likely an irregular movement is to happen in the individual block. Then again, we can say that there are abnormal activities in t continuous frames if a higher worth exists in the minimum-distance matrix. Subsequently, we locate the most noteworthy incentive in the minimum-distance matrix as the frame illustrative feature value. In the event if the most elevated value of the minimum-distance matrix is bigger than the edge, we group the current frame as abnormal.

Pixel level identification of abnormal activities. Once a frame is recognized as abnormal, we look at the estimation of the minimum-distance matrix of each mega block with the edge value.

4. Modelling and Implementation Details

4.1 Design Diagrams

4.1.1 Control Flow Diagram

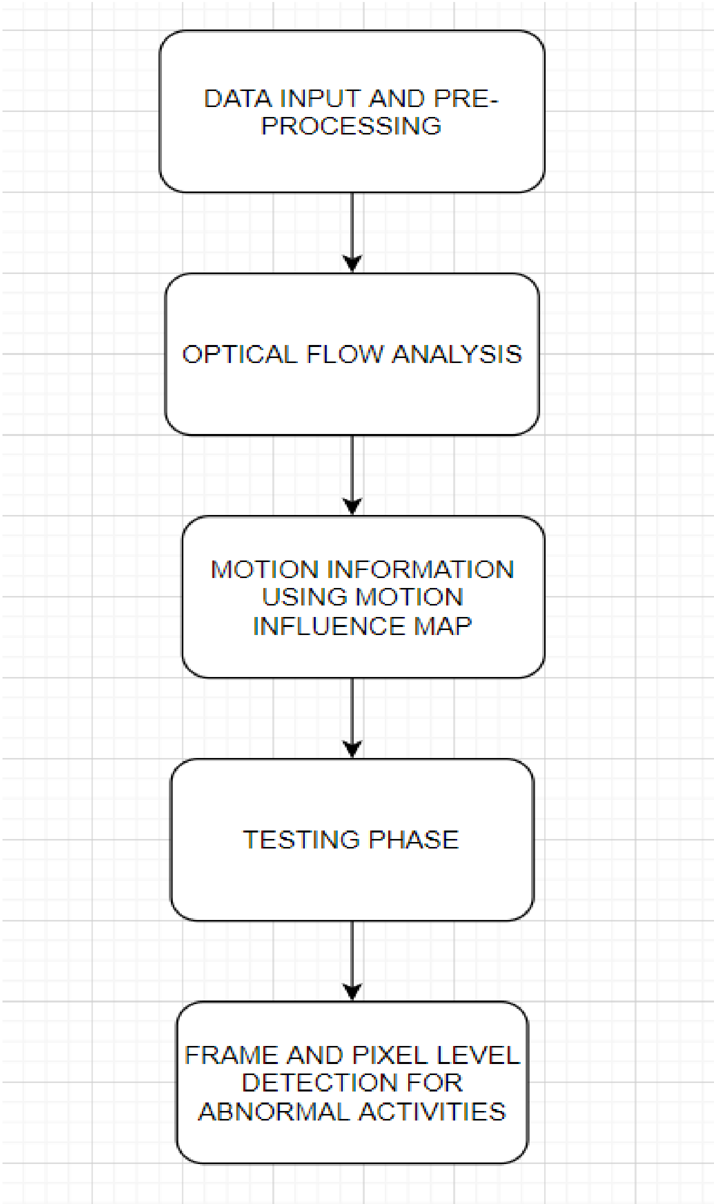
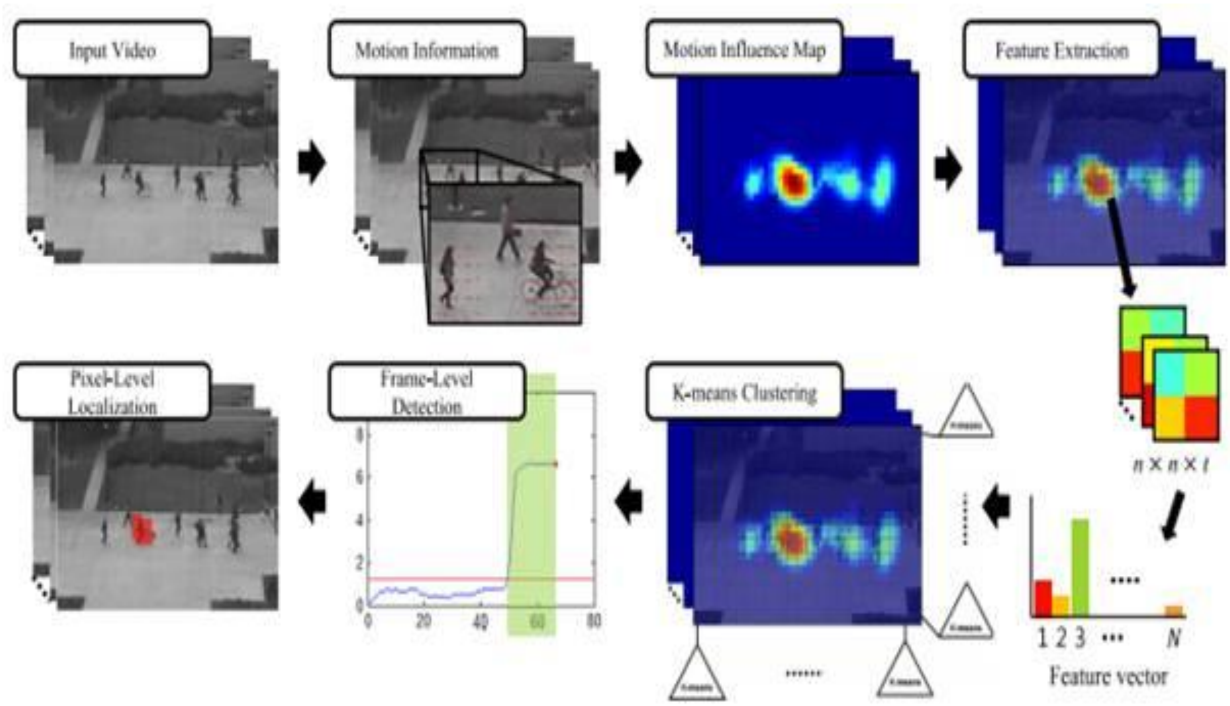
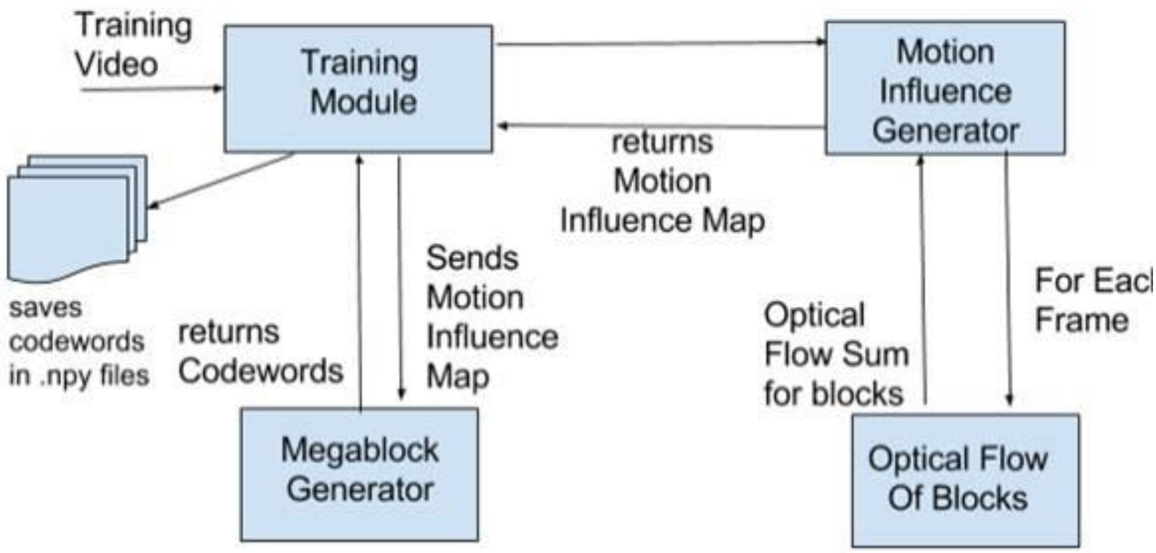


Fig 7: Control flow diagram

4.1.2 Sequence Diagram



(a)



(b)

Fig 8(a,b): Sequence diagram

5. Testing (Focus on Quality of Robustness and Testing)

Software quality Robustness and Testing is a non-functional requirement for a software program which is not called up by the customer's contract, but nevertheless is a desirable requirement which enhances the quality of the software program.

- **Completeness:** Presence of all constituent parts, with each part fully developed.
- **Conciseness:** Minimization of excessive or redundant information or processing.
- **Portability:** Ability to be run well and easily on multiple computer configurations. Portability can mean both between different hardware
- **Testability:** Disposition to support acceptance criteria and evaluation of performance. Such a characteristic must be built-in during the design phase if the product is to be easily testable; a complex design leads to poor testability.
- **Usability:** Convenience and practicality of use. This is affected by such things as the human-computer interface.
- **Reliability:** Ability to be expected to perform its intended functions satisfactorily. This implies a time factor in that a reliable product is expected to perform correctly over a period of time.
- **Efficiency:** Fulfillment of purpose without waste of resources, such as memory, space and processor utilization, network bandwidth, time, etc.
- **Security:** Ability to protect data against unauthorized access and to withstand malicious or inadvertent interference with its operations.

5.1 Testing Plan

We started the testing phase by first carrying out unit testing in the entire project and then integrating testing was applied while integrating the modules so that there are no bugs while using functions of other modules while calling from some other module

Type of Test	Will Test Be Performed?	Comments/Explanations	Software Component
Unit	Yes ▪ No	Every component has been individually tested before being used as an integrated module	optFlowofblocks.py,motionInfluenceGenerator.py, training.py
Integration	Yes ▪ No	After all individual modules were working, the integrated code was tested	Test.py

Table 1: testing

5.2 Limitations of The Solution

The proposed method has a limitation when there is a strong perspective distortion in the input video as the motion influence map is built based on the motion direction and magnitude of the moving objects. However, the main focus of this work is to detect unusual activities within a crowded scene, for which the cameras usually cover a wide area, resulting in small objects being present in the scene without significant perspective changes. Also, our experiments were limited to a fixed viewpoint, and there is a limitation in the applicability of the approach for surveillance cameras with pan, zoom, or tilt functionality. At this moment the proposed method deals only with static cameras. However, it can be easily extended to PTZ cameras using localization results. In this regard, we assume that there are no large scaling changes or pans, tilt, or zooms in a scene, which may be a potential limitation of our method.

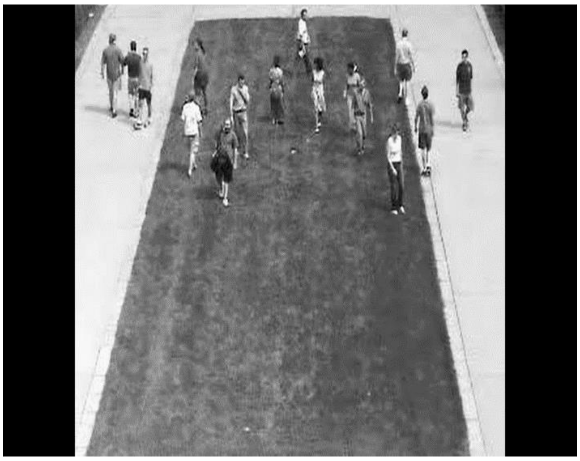
6. Findings, Conclusion, and Future Work

6.1 Findings

We partitioned each frame into 8x8 nonoverlapping blocks, and set the threshold to the maximum feature value in the motion influence maps of the training images.

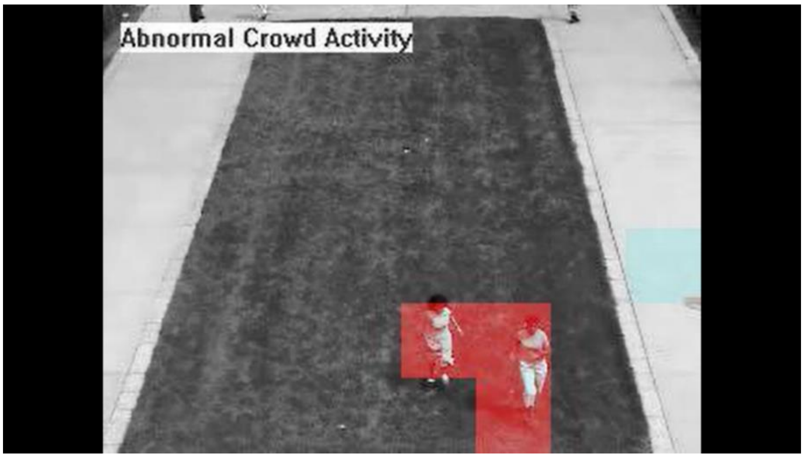
Snapshots in Scene 1 and scene 2 shows the normal and abnormal activities being detected in the testing phase. Scene 1 (a) shows normal people just walking in the garden i.e is considered as normal activity whereas in scene (b) and (c) people randomly starts running around that shows the abnormal activity and people running are being detected which is shown in the below figures.

Scene 1:



(a) : Normal Activity

(b): Abnormal Activity being detected



(b): Abnormal Activity being detected

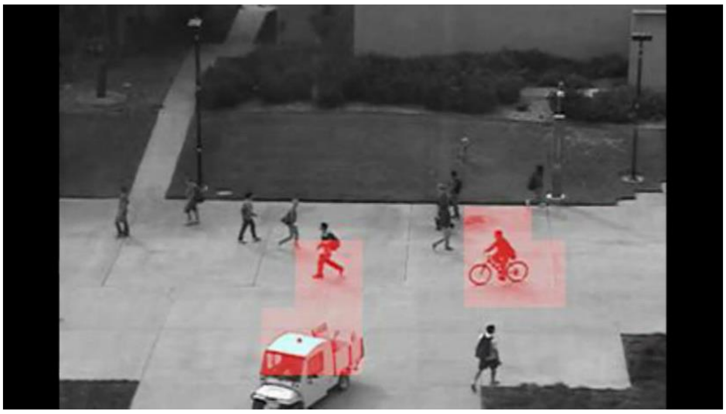
Fig 9(a,b,c): Abnormal activity detection scene 1

Scene 2 people walking on the road. In scene 2(a) as we can see abnormal activity is being detected one person is moving on the cycle and another on skateboard while others are walking on the road. In scene 2(b) we can see abnormal activity is being detected one person is moving in the opposite direction as compared to others. In scene 2(c) we can see that a car and a bicycle is being detected as abnormal activity.

Scene 2:

(a): Abnormal Activity being detected

(b): Abnormal Activity being detected



(c): Abnormal Activity being detected

Fig 10(a,b,c):Abnormal activity detection scene 2

In the below figure we can see the motion generator map with respect to the scene where abnormal activity is being detected. This helps in marking the abnormal activity in the scene as compared to normal activity.

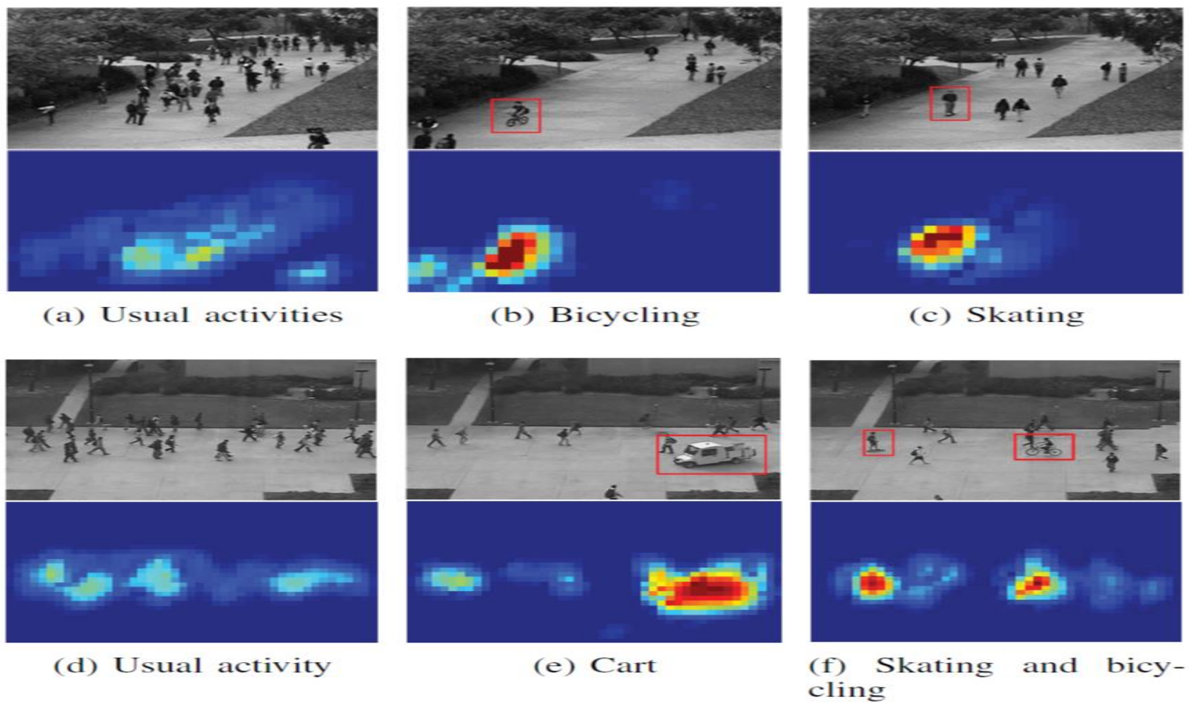


Fig 11: Sample frames of various unusual activities

After the testing phase we got these results: -

Accuracy	93.37%
Precision	83.19%
Recall	98.34%
F-score	0.861553

Table 2: Accuracy, Precision, Recall, F-score

Finally, after testing the model which we have created on the dataset which contains both normal and abnormal activities we have got the accuracy of 93.37% i.e. our model is working efficiently.

6.1.1 Snapshots

```
Epoch 96/100
208/208 [==============================] - 0s 526us/step - loss: 0.0072 - acc: 0.9952 - val_loss: 0.3888 - val_acc: 0.
8667
Epoch 97/100
208/208 [==============================] - 0s 531us/step - loss: 0.0072 - acc: 0.9952 - val_loss: 0.3920 - val_acc: 0.
8667
Epoch 98/100
208/208 [==============================] - 0s 531us/step - loss: 0.0072 - acc: 0.9952 - val_loss: 0.4079 - val_acc: 0.
8667
Epoch 99/100
208/208 [==============================] - 0s 523us/step - loss: 0.0073 - acc: 0.9952 - val_loss: 0.4168 - val_acc: 0.
8667
Epoch 100/100
208/208 [==============================] - 0s 514us/step - loss: 0.0072 - acc: 0.9952 - val_loss: 0.4031 - val_acc: 0.
8556

<keras.callbacks.History at 0x7f528d0f7080>
```

Fig12 : Snapshot 1 (Training losses)

```
  [0.00000000e+00 0.00000000e+00 0.00000000e+00 ... 3.45010512e-27
   0.00000000e+00 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 6.20573201e-06 ... 1.14387712e-02
   3.74954902e-02 0.00000000e+00]]

 ...

 [[7.30475034e-36 2.44276881e-35 3.38639164e-14 ... 8.94276840e-13
   6.81854045e-17 7.95155603e-40]
  [0.00000000e+00 0.00000000e+00 1.28431701e-16 ... 9.63082840e-08
   1.19546531e-02 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 7.08354193e-28 ... 2.38064451e-13
   0.00000000e+00 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 0.00000000e+00 ... 6.91979172e-16
   0.00000000e+00 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 2.81270819e-33 ... 0.00000000e+00
   0.00000000e+00 0.00000000e+00]]

 [[0.00000000e+00 0.00000000e+00 8.84169439e-22 ... 3.40835554e-11
   4.33531677e-04 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 4.02187793e-39 ... 5.93513891e-29
   0.00000000e+00 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 0.00000000e+00 ... 1.26304113e-34
   0.00000000e+00 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 0.00000000e+00 ... 0.00000000e+00
   0.00000000e+00 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 7.39191316e-28 ... 8.27649985e-25
   2.87626756e-28 0.00000000e+00]]

 [[0.00000000e+00 0.00000000e+00 1.40154255e-25 ... 7.22323101e-15
   7.01336103e-06 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 9.13739645e-39 ... 1.02294788e-42
   5.07362390e-39 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 0.00000000e+00 ... 0.00000000e+00
   0.00000000e+00 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 0.00000000e+00 ... 0.00000000e+00
   0.00000000e+00 0.00000000e+00]
  [0.00000000e+00 0.00000000e+00 0.00000000e+00 ... 0.00000000e+00
   0.00000000e+00 0.00000000e+00]]]]
Done
```

Fig13: Snapshot 2 (Bag of codewords formed after training for normal activities)

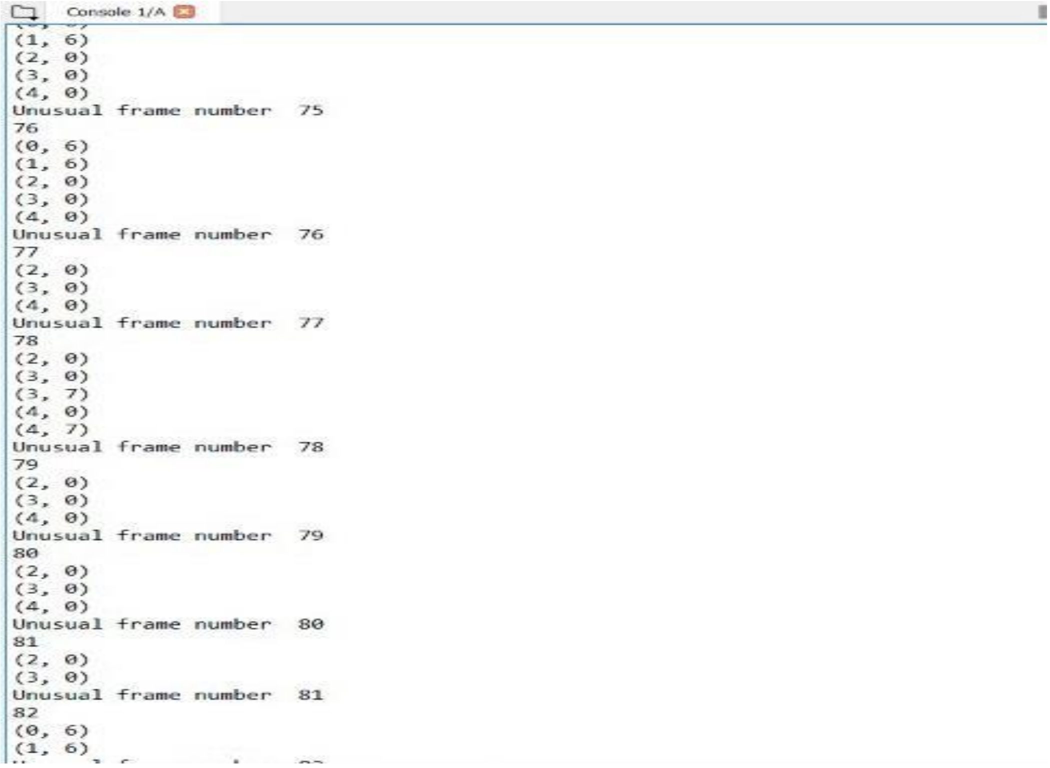


Fig14: Snapshot 3 (Unusual frames detected)

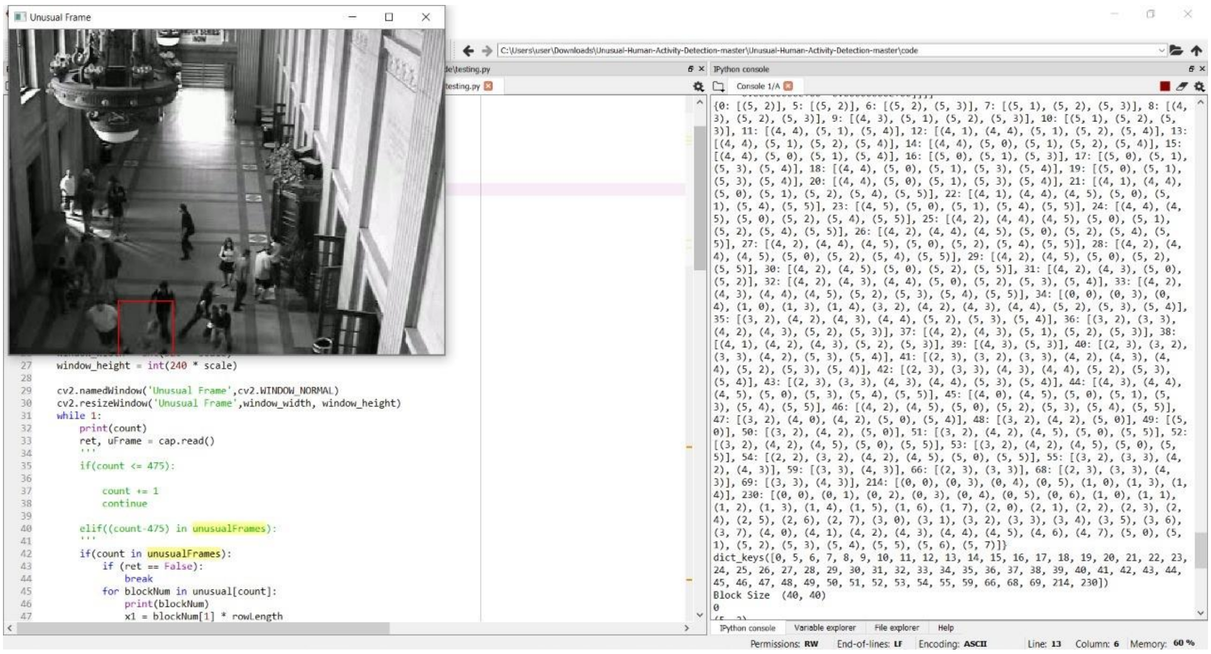


Fig15: Snapshot 4 (Display of result)

6.2 Conclusion

We have presented an abnormal event detection method via motion influence map. This approach directly learns motion influence characteristics, which increase the testing speed hundreds of times without compromising effectiveness. Our method achieves state-of-the-art results in several datasets. It is related to but differ largely from traditional subspace clustering.

In a time where surveillance cameras are being used everywhere, effectively checking it for any abnormal event would be a bottleneck. Thus, a fast and intelligent method to check theses surveillance cameras is at most required. It would help in cutting down a lot of work to be done by people struggling to monitor it and would help it taking faster actions during those situations by integrating these with alarms and other important actions like informing the police or calling an ambulance.

Since it achieves a frame rate of 100fpm, frames can be analysed at a decent rate and thus can be used in surveillance cameras to detect abnormalities automatically. Based on the signal of this system, alarms and other actions can be controlled.

6.3 Future Work

- Our future work will be to extend the motion influence map framework to other video applications. This can be extended to detect various other kind of abnormal event that are usually encountered. By doing this a system which can detect any abnormality will be build which can deployed in various environments just encouraging portability. The algorithm should be carefully analysed for areas where parallel processing is possible and suitably the algorithm should be tweaked which can help in achieving better accuracy.
- The model can be extended by extracting various other features that would help in detection of other abnormalities. For example: Detection on knife or guns or anything suspicious objects.
- The model that can used to fulfil the need of the current open challenge by integrating it with the audio interface. For example: A woman running and shouting for help in an area could be detected as unusual.
- Model can also be integrated with the facial recognition system and can be helped in criminal identification.

REFERENCES

[1] Attard, L., & Farrugia, R. A. (2011, April). Vision based surveillance system. In *2011 IEEE EUROCON-International Conference on Computer as a Tool* (pp. 1-4). IEEE.

[2] Khan, K., Albattah, W., Khan, R. U., Qamar, A. M., & Nayab, D. (2020). Advances and trends in real time visual crowd analysis. *Sensors*, *20*(18), 5073.

[3] Gnouma, M., Ejbali, R., & Zaied, M. (2018). Abnormal events' detection in crowded scenes. *Multimedia Tools and Applications*, *77*(19), 24843-24864.

[4] Rodriguez, M. D., & Shah, M. (2007, September). Detecting and segmenting humans in crowded scenes. In *Proceedings of the 15th ACM international conference on Multimedia* (pp. 353-356).

[5] Sultani, W., Chen, C., & Shah, M. (2018). Real-world anomaly detection in surveillance videos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 6479-6488).

[6] Lee, D. G., Suk, H. I., Park, S. K., & Lee, S. W. (2015). Motion influence map for unusual human activity detection and localization in crowded scenes. IEEE Transactions on Circuits and Systems for Video Technology, 25(10), 1612-1623.

[7] Wang, Y., Zhang, Q., & Li, B. (2016, March). Efficient unsupervised abnormal crowd activity detection based on a spatiotemporal saliency detector. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1-9). IEEE.

[8] Huang, Y., Liu, Z., Jiang, M., Yu, X., & Ding, X. (2019). Cost-effective vehicle type recognition in surveillance images with deep active learning and web data. *IEEE Transactions on Intelligent Transportation Systems*, *21*(1), 79-86.

[9] Boult, T. E., Micheals, R. J., Gao, X., & Eckmann, M. (2001). Into the woods: Visual surveillance of noncooperative and camouflaged targets in complex outdoor settings. *Proceedings of the IEEE*, *89*(10), 1382-1402.

[10] Mehran, R., Oyama, A., & Shah, M. (2009, June). Abnormal crowd behavior detection using social force model. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 935-942). IEEE.

[11] Hara, K., Omori, T., & Ueno, R. (2002, September). Detection of unusual human behavior in intelligent house. In *Proceedings of the 12th IEEE workshop on Neural Networks for Signal processing* (pp. 697-706). IEEE.

[12] Khan, Z. A., & Sohn, W. (2011). Abnormal human activity recognition system based on R-transform and kernel discriminant technique for elderly home care. *IEEE Transactions on Consumer Electronics*, *57*(4), 1843-1850.

[13] Chen, L., Li, Y., & Liu, Y. (2020, August). Human body gesture recognition method based on deep learning. In *2020 Chinese Control And Decision Conference (CCDC)* (pp. 587-591). IEEE.

[14] Miao, Z., Zou, S., Li, Y., Zhang, X., Wang, J., & He, M. (2016). Intelligent video surveillance system based on moving object detection and tracking. *DEStech Transactions on Engineering and Technology Research*, (iect).

[15] Shidik, G. F., Noersasongko, E., Nugraha, A., Andono, P. N., Jumanto, J., & Kusuma, E. J. (2019). A Systematic Review of Intelligence Video Surveillance: Trends, Techniques, Frameworks, and Datasets. *IEEE Access*, *7*, 170457-170473.

[16] Konda, K. R., Tefera, Y. T., Conci, N., & De Natale, F. G. (2017, June). Real-time moving object detection and segmentation in H. 264 video streams. In *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)* (pp. 1-6). IEEE.

[17] Guo, H., Liang, Y., Yu, Z., & Liu, Z. (2010, June). Implementation and analysis of moving objects detection in Video Surveillance. In *The 2010 IEEE International Conference on Information and Automation* (pp. 154-158). IEEE.

[18] Mohite, A. S., Sangale, D. K., Oza, P. R., Parekar, T. D., & Navale, M. P. (2020). Unusual Human Activity Detection Using Opencv Python with Machine Learning. *CLIO An Annual Interdisciplinary Journal of History*, *6*(3), 183-187.

[19] Balaji, S. R., & Karthikeyan, S. (2017, January). A survey on moving object tracking using image processing. In *2017 11th international conference on intelligent systems and control (ISCO)* (pp. 469-474). IEEE.

[20] Dang, L. M., Min, K., Wang, H., Piran, M. J., Lee, C. H., & Moon, H. (2020). Sensor-based and vision-based human activity recognition: A comprehensive survey. *Pattern Recognition*, *108*, 107561.

[21] Khan, M. U. K., Park, H. S., & Kyung, C. M. (2018). Rejecting motion outliers for efficient crowd anomaly detection. *IEEE Transactions on Information Forensics and Security*, *14*(2), 541-556.

[22] Nanaware, V. S., Nerkar, M. H., & Patil, C. M. (2017, September). A review of the detection methodologies of multiple human tracking & action recognition in a real time video surveillance. In *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)* (pp. 2484-2489). IEEE.

[23] Momin, B. F., & Fanase, V. M. (2015, April). Detection and localization of abnormal activities in video surveillance system. In *2015 International Conference on Communications and Signal Processing (ICCSP)* (pp. 0277-0280). IEEE.

[24] Ramprasadi, N., Shah, P., & Vyas, D. (2020, May). Hybrid Approach For Real Time Crowd Activity Identification Using Segmentation. In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 390-394). IEEE.

[25] Wang, L., & Dong, M. (2012, September). Real-time detection of abnormal crowd behavior using a matrix approximation-based approach. In *2012 19th IEEE International Conference on Image Processing* (pp. 2701-2704). IEEE.

[26] Bertini, M., Del Bimbo, A., & Seidenari, L. (2012). Multi-scale and real-time non-parametric approach for anomaly detection and localization. *Computer Vision and Image Understanding*, *116*(3), 320-329.

[27] Jhapate, A. K., Malviya, S., & Jhapate, M. (2020, February). Unusual Crowd Activity Detection using OpenCV and Motion Influence Map. In *2nd International Conference on Data, Engineering and Applications (IDEA)* (pp. 1-6). IEEE.

[28] Jobard, B., & Lefer, W. (1997, October). The motion map: efficient computation of steady flow animations. In *Proceedings. Visualization'97 (Cat. No. 97CB36155)* (pp. 323-328). IEEE.

[29] Yao, L., Liu, Y., & Huang, S. (2016). Spatio-temporal information for human action recognition. *EURASIP Journal on Image and Video Processing*, *2016*(1), 39.

[30] Rasheed, N., Khan, S. A., & Khalid, A. (2014, May). Tracking and abnormal behavior detection in video surveillance using optical flow and neural networks. In *2014 28th International Conference on Advanced Information Networking and Applications Workshops* (pp. 61-66). IEEE.