I.J. Image, Graphics and Signal Processing, **2022, 6, 15-25**
Published Online on December 8, 2022 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2022.06.02

Motion Pattern Based Anomalous Pedestrian Activity Detection

Kamal Omprakash Hajari*
Yeshwantrao Chavan College of Engineering, Rashtrasant Tukadoji Maharaj Nagpur University, Department of Information Technology, Maharashtra, India – 441110
E-mail: kamalhajari123@gmail.com
ORCID iD: https://orcid.org/0000-0002-4959-8117
*Corresponding Author

Ujwalla Haridas Gawande
Yeshwantrao Chavan College of Engineering, Rashtrasant Tukadoji Maharaj Nagpur University, Department of Information Technology, Maharashtra, India - 441110
E-mail: ujwallgawande@yahoo.co.in
ORCID iD: https://orcid.org/0000-0002-7059-0937

Yogesh Golhar
St. Vincent Palloti College of Engineering and Technology, Rashtrasant Tukadoji Maharaj Nagpur University, Department of Computer Engineering, Maharashtra, India - 441110
E-mail: yogesh.golhar28@gmail.com
ORCID iD: https://orcid.org/0000-0002-6817-3552

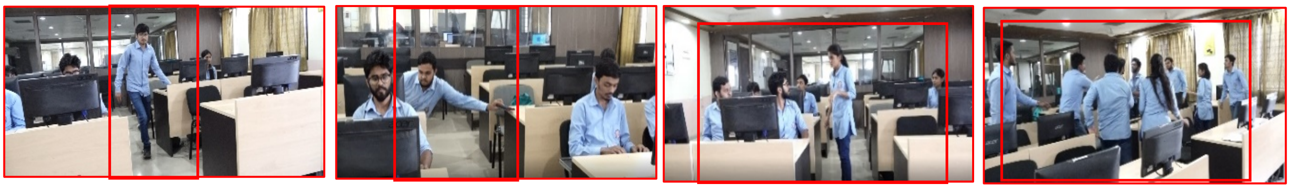
Received: 17 April, 2022; Revised: 08 June, 2022; Accepted: 26 August, 2022; Published: 08 December, 2022

Abstract: In this paper, an efficient technique for anomalous pedestrian activity detection in the academic institution is proposed. At the pixel and block levels, the proposed method elicits motion components that accurately represent pedestrian action, velocity, and direction, as well as along a frame. We also adopted these motion features to detect anomalous actions. The detection of anomalous behavior in academic environments is not available at the moment. Similarly, the existing method produces a high number of false positives. An anomaly detection dataset and a newly designed proposed student behavior database were used to validate the proposed framework. A significant improvement in anomalous activity recognition has been demonstrated in experimental results. Based on motion features, the proposed method reduces false positives by 3% and increases true positives by 5%. A discussion of future research directions concludes the paper.

Index Terms: Artificial Intelligence, Computer vision, Pedestrian dataset, Tracking, Detection, Motion Pattern, Anomalous activity.

1. Introduction

Surveillance cameras are now used to keep an eye on both public and private notable areas. Monitoring the conduct of pedestrians under human supervision is a challenging undertaking. It is difficult and time-consuming to identify and separate anomalous behavior patterns from video in real-time with the conventional approach. We need a system that can identify anomalous actions in real-time due to system restrictions. Recognizing human movement and activity in video frames has drawn the attention of numerous academics and experts in this field in recent years [1, 2, 3]. In recent years, expert's attention has shifted towards spotting anomalous activities in densely populated areas. Pedestrians cannot be detected and tracked in high-density areas due to factors like partial or total occlusion of objects, changes in object size, and variations in ambient lighting. The use of texture-based information for identifying unusual behaviors in crowded situations has been attempted by numerous authors. Example techniques include temporal gradients [4], dynamic texture characteristics [5], and spatio-temporal frequency features [6, 7]. Optical flows have also been studied, which can be used in video frames to identify motion features directly. Examples include motion pattern clusters [8, 9, 10], force map generation for congested area behavior prediction [12], heat map generation for motion pattern map generation [8], spatial motion feature [11], trajectory-based object motion analysis [14], motion of individual object detection using an optical flow of motion directions [13], and isolated object motion pattern intensity based histography [13]. The motion flow pattern strategy has been shown to be useful in prior research, and we think there is still an opportunity for development in the current approaches. It is crucial to include information on objects of various sizes, their motions, speeds, and interactions between frames. By carefully examining the information related to this motion, we can improve our performance. The information concerning this movement can help us perform better if we take it into account. Regardless of the size and orientation of the object, the proposed methodology extracts the movement of the moving pedestrian.



(a) (b)

Fig. 1. Example of academic activity student anomalous behavior. (a) Mobile phone stealing (Local region). (b) Dispute in the lab. (Global region).

We outline a brand-new motion component that effectively captures the motion patterns of objects in a congested space. In this method, distinct activity areas in video frames into two groups: local and global. Both the region of extraordinary activity and the zone of regular activity are shown in Fig.1. A local area is a discrete active area inside an abnormal area. A global region, on the other hand, denotes the area where odd behaviors are seen. Different methods of determining the region's extraordinary activity have been introduced in the literature. In [15] the author proposed a social force map-based method for spotting global aberrant behavior. A particle grid was used to map the optical flow field, and the forces exerted by the particles were determined. In [11] the author described a method for spotting fake local motions in a scene. In order to locate abnormalities close by, they created a subspace feature map employing optical flow features of various sizes. A real-time monitoring system needs a comprehensive framework that can identify odd student's behavior. We put forth a special method for spotting localized and widespread aberrant student conduct in a crowded setting. We evaluated the proposed approach using our own student behavior dataset as well as other benchmark datasets, such as the University of Minnesota anomaly detection datasets. The proposed contribution has the following objectives:

1. The main objective of the motion-based unusual activity recognition systems is to make surveillance systems smarter and more intelligent.
2. To develop the active video surveillance system that can work efficiently in the real-time environment and it can be used as active and real-time automated medium of security.
3. To detect anomalous or abnormal activities in videos to avoid future happening or to alert whenever any type of mis- happening occur.
4. To detect intruders around the perimeter of the main building.

The proposed contributions are as follows:

1. To solve existing state-of-the-art database concerns such as size and illumination variance in pedestrian images, we present the motion pattern based unusual activity detection framework.
2. We propose a student activity dataset. In which we have recorded the student's normal and anomalous activities.
3. Within the framework of the proposed pedestrian dataset for academic settings, we conduct a comprehensive review of previous work and compare existing techniques.

The remainder of the contribution is structured as follows: In Section 2, the current contribution in this field is described. In Section 3, we suggest a motion feature-based anomalous activity detection system. A novel pedestrian, a database of common behaviors created in an academic environment, and experimental findings are covered in Section 4. A research direction is given in the final section.

2. Literature Survey

The most recent and pertinent datasets on pedestrians are described in this section. We also go through the limitations of advanced deep learning algorithms for detecting, following, and identifying anomalous conduct in pedestrians.

2.1. Anomalous activity detection recent methodology

The detection of unexpected activity has recently attracted the attention of researchers in smart video surveillance systems. The authors of [17] discuss how challenging it is to interpret and model behavior from surveillance footage.

Violations were found utilizing likelihood ratio analysis and categories of pedestrian legal action in an unsupervised learning framework. The author of [18] offered a method for locating anomalies in a spatial and temporal context. They showed an immediate event for a specific object in a scene that includes the object's position, movement, direction, and velocity. They combine three partitioned atomic events to describe legitimate occurrences. The aforementioned strategies are not appropriate for situations when it is difficult to detect and monitor a single pedestrian in high-density locations, under various lighting conditions, and at various scales. Other more recent studies use motion flow and directional data to track and recognize pedestrians. The authors in [19] use the Kanade Lucas-Tomasi (KLT) technique, in which things are represented utilizing the motion of the objects. It employs two distinct categories of historical and self-history descriptors, as well as the histories of surrounding objects, to detect anomalies in a scene [10]. The author of [20] described a method for manually counting the number of persons in a photograph. Along with foreground relationships, they also used an optical-flow motion pattern. They calculated the crowd's exponential distribution patterns as well as the dynamic energy of using optical flow to distinguish between running and walking motions. Understanding and simulating densely populated people behavior has been the focus of several academics [21, 22, 23, 24]. By modelling the behavior of the crowd, several techniques were employed to find abnormal activity around the world. The social force model was employed by the author in [22] to describe crowd behavior. This technique locates and detects the object in a scene by utilizing the optical flow of moving objects [25].

The social force, which was calculated via latent Dirichlet allocation, is used to distribute normal activities. The authors of [26] examine social behavior and its behavioral using interactive energy potentials. The moving objects in a scene are represented as spatiotemporal interest points [27], and a KLT tracker is used to keep track of each object in a frame [19] [33]. On the other hand, other research teams have focused on identifying regional aberrant activity. The phrase "quantifiable" is used by the author to describe how uncommon it is to choose uncorrelated motions from a geographical environment in [11]. They identified the index in numerous channels that were travelling in diverse directions and at various rates. Nearly all the anomalous behavior in the area could be detected using the linear map that was provided. The author of [8] used motion intensities to build a motion heat map and compared it to local motion fluctuations. The authors of [5] use atypical behaviors that are localized in crowded settings utilizing spatial cues. Once more, crowd behavior analysis has been done using spatio-temporal cuboids that were recovered from optical flow or gradient pattern characteristics [29, 34, 35, 36, 37]. The aforementioned techniques have been found to be effective in research, although they are often only capable of spotting odd activity in a local or global setting. In present state-of-the-art view, depicting pedestrian activity in a high-density scene can be accomplished by considering the motion flows pattern, changing item sizes, and interactions between close-by objects in a frame [38, 39, 40]. This can improve performance in detecting anomalous activity. In this study, we offer a powerful technique for capturing motion data and employing a motion feature to identify both local and global aberrant behaviors [40, 41, 42].

2.2. Pedestrian Dataset

We outline pedestrian datasets used by scientists to identify anomalous activities and detect and track pedestrians. First, there are 350 000 annotated bounding boxes in the Caltech dataset that represent 2,300 distinct pedestrians. Using the camera installed on the car and a city road, this dataset was produced [3]. The second dataset is the well-known pedestrian dataset from MIT, which includes examples of high-quality pedestrian images. There are 709 different pedestrians present. The variety of posture images captured on city streets [4], whether in front view or back view, is somewhat constrained. Third, Daimler, this dataset documents pedestrian activity in an urban setting during daylight hours using cameras mounted on moving vehicles. The dataset consists of ground truth images, floating disparity map files, annotated labeled bounding boxes, and pedestrian tracking features. 15560 pedestrian images and 6744 annotated pedestrian images make up the training set. 56,492 annotated images and 21,790 images of pedestrians make up the test set [5]. The ATCI dataset, a database of pedestrians collected by a typical car's rear-view camera, is used to evaluate pedestrian detection in parking lots, metropolitan settings, city streets, and private as well as public lanes. The data collection consists of 200,000 indicated pedestrian bounding boxes and 250 video clips, each lasting 76 minutes and taken under daylight conditions with contrast weather conditions [6]. The traffic scene is observed from the inside of the car using the ETH dataset. Over the car, the actions of pedestrians are seen. The dataset can be used in an urban environment for mobile platform-based pedestrian tracking and recognition. The dataset includes both on-road vehicles and pedestrians [7]. The TUD-Brussels dataset was produced in an urban setting on a mobile device. The vehicle's front side was equipped with a camera that was used to capture the activity of crowded metropolitan streets. It can be used in urban driving safety scenarios [8]. The INRIA dataset is one of the most ethereal pedestrian detection datasets. It includes human activity along with a moving camera and intricate background scenes with different variations in posture, appearance, clothing, background, lighting, contrast, etc. [9]. Static objects with varying angles and positions can be found in an urban scene in the PASCAL Visual Object Classes (VOC) 2017 and 2007 collection. The objective of this dataset was to identify visual item types in practical situations. There are 20 distinct categories in this collection, including those for people, cars, trees, and road signs [10]. The MS COCO 2018 dataset was used to create the Common Object in Context [11].

Recently, the 2018 dataset was used with an emphasis on stimulus item detection to identify various things in the context. The annotations provide instances of various items related to 91 distinct human segmentation categories and 80 distinct object categories. There are five picture labels per sample image and key point annotations for pedestrian examples. The COCO 2018 dataset tasks include (1) real-scene item recognition with a segmentation mask, (2) panoptic segmentation, (3) pedestrian key point evaluation, and (4) Dense Pose estimation in a congested scene [12]. The Mapillary Vistas Research dataset is used to segment street pictures [13]. With the use of panoramic segmentation, which successfully combines the ideas of semantic and instance segmentation, pedestrians and other non-living categories are solved. Table 1 compares pedestrian databases and the reasons for their video surveillance. We've also added our suggested dataset, which is described in the section after this one. The usage, size, environment, label, and annotation of the dataset all play a role in the link. These specifics are needed to confirm that the object detection and tracking algorithms are accurate.

Table 1. Comparison of benchmark pedestrian dataset.

Dataset	Dataset size	Annotation	Environment	Year	Ref.	Issues and Challenges
Caltech	250000 frames	2300 unique	City street	2012	[3]	Only urban roads are captured.
MIT	709 unique pedestrians	No annotated pedestrian	Day light scenario	2000, 2005	[4]	Missing annotation not allow user to verify different techniques.
Daimler	15,560 unique pedestrians	Ground truth with bounding boxes.	City street	2016	[5]	Only urban roads are captured.
GM-ATCI	Video clips:250	Annotated pedestrian bounding boxes: 200K	Day and complex Weather and lighting	2015	[6]	Only urban roads are captured. Side view of road not captured.
ETH	Videos	Annotated cars and pedestrians	City street	2010	[7]	Small size dataset. Limited scenarios cover.
TUD Brussels	1092 frames	Pedestrian Annotation 1776	City street	2009	[8]	Only urban roads are captured.
INRIA	498 images	Manual Annotations	City street	2005	[9]	Only urban roads are captured.
PASCAL VOC 2012	11,530 images, 20 objects classes	ROI Annotated 27,450	City street	2012	[10]	Only urban roads are captured.
MS COCO 2017	328,124 images	Segmented people object	City street	2017	[11]	Only urban roads are captured.
MS COCO 2015	328,124 images	Segmented people object	City street	2015	[12]	Only urban roads are captured.
Mapillary Vistas	152 objects , 25,000 images	Pixel accurate instance Pedestrian	City street	2017	[13]	Only urban roads are captured. Side view of road not captured.

3. Proposed Methodology

The proposed method accurately depicts the moving object features in dense locations by exploiting data on pedestrian motion. We noticed the quick and deliberate movement of an unusual pedestrian activity in both the local and global parts of the image. The general design of the proposed method is depicted in Fig.2. Each frame is split into blocks, from which motion data is then extracted to create motion characteristics at the pixel and block levels.

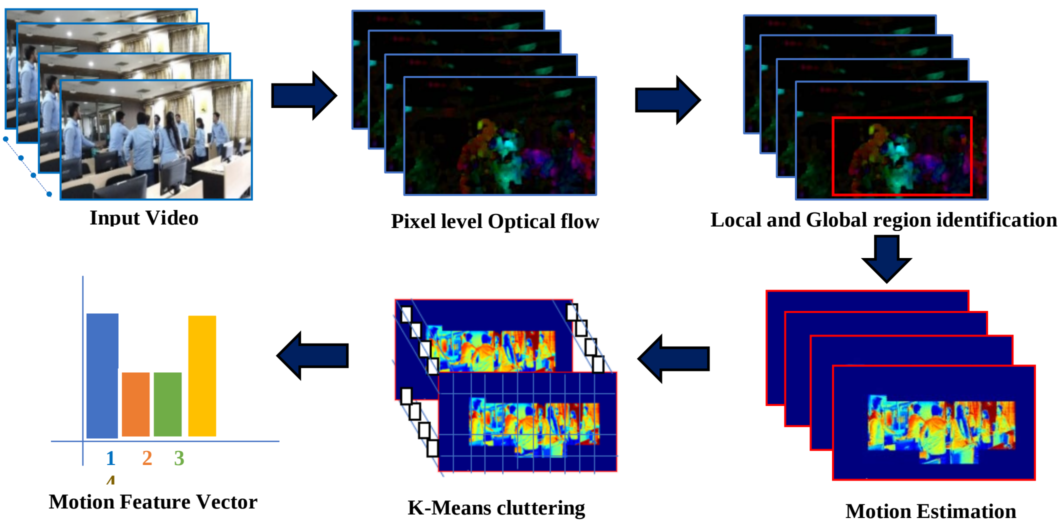


Fig. 2. Proposed framework architecture for motion feature extraction using Pixel-level Optical Flow.

The following steps comprise the motion feature extraction process:

1. The activities of moving objects are segmented at the pixel and block levels and computed progressively throughout a set of frames.
2. First, the motion information is extracted at pixel level using the optical flow method. The motion feature represented by (u, v) that encodes the displacement in x and y direction between pixels in an 8-neighborhood. i.e. (left, right, up-left, up-right, down, down-left, and down-right). The optical flow is computed for each pixel $I(x, y, t)$. It changes its location by (dx, dy) at time dt. It is represented by Eq. 1.

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \tag{1}$$

3. Next, the following equation is obtained by taking the Taylor series approximation of the right-hand side, eliminating common phrases, and dividing by dt:

$$f_x u + f_y v + f_t = 0 \tag{2}$$

Were,

$$f_x = \frac{\partial f}{\partial x} \text{ and } f_y = \frac{\partial f}{\partial y}$$

$$u = \frac{\partial x}{dt} \text{ and } v = \frac{\partial y}{dt}$$

Here, f_x and f_y are the gradient and f_t is the gradient over time. (u, v) are the unknown term computed using the Lucas-Kanade method and it is represented by Eq. 3.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_i {fx_i}^2 & \sum_i fx_i\, fy_i \\ \sum_i fx_i\, fy_i & \sum_i {fy_i}^2 \end{bmatrix} \begin{bmatrix} -\sum_i fx_i\, ft_i \\ -\sum_i fy_i\, ft_i \end{bmatrix}$$

4. Different orientations of the motion characteristics are retrieved. A single feature matrix that embodies both spatial and temporal properties is produced after the integration of motion information.
5. To categorise the activity, k-means clustering is done for each zone to pinpoint the global and local region.
6. Frame-level clustering produces different clustered motion sequence that is represented as features. Extracted features contain the motion of an item in each frame.
7. Using pixel-level localization, we may pinpoint the precise position of the unexpected behaviour after classifying a frame as unusual. The procedure will go on until the entire video sequence has been produced.

A technique that involves going through each frame of a video stream and looking for anomalous human activity. Each pixel's optical flow in Eq. 3 is first extracted using the proposed method.

$$Block_i = \frac{1}{J} \sum_{j=0}^{n} f(x, y)_i^j \tag{3}$$

Where, $Block_i$ represented by the i^{th} position in block of optical flow, each pixel in a frame size is denoted by J, i^{th} block j^{th} pixel of each individual frame of optical flow is denoted by $f(x, y)_i^j$.

Furthermore, each block threshold is computed, which is represented by Th_d using optical flow motion pattern denoted by $Block_i$ and it is having the width Wd computed using the Eq. 4.

$$Th_d \leftarrow Block_i \; x \; Wd \tag{4}$$

If the $EucliD_{(i,j)}$ less than the threshold then the output is 1 and in other condition it is set to 0. The angle between the motion feature vector represented by Θ_{ij}.

$$\Theta_{ij} \begin{cases} 1 & EucliD_{(i,j)} < Th_d \\ 0 & \text{Otherwise} \end{cases}$$

The motion feature vector is computed using $MotF\ (\Theta B_i)$, $EucliD_{(i,j)}$, and $Block_i$ represented by Eq. 5.

$$MotF = \frac{MotF\ (\Theta B_i) + EucliD_{(i,j)}}{Block_i} \tag{5}$$

wherein MF is the motion vector of the block. The Euclidean distance between ith and jth block is represented by $EucliD_{(i,j)}$. The algorithm that follows provides a basic description of motion feature extraction. Additionally, the motion region in a frame was clustered at the frame level, with each cluster's optical pixel flow in a distinct direction being considered the feature vector. The likelihood of unexpected behavior in the matching block decreases as the distance between them grows closer together. The classification of anomalous activities in additional frames is possible if a higher value of distance is found. So, the current scene is identified as an uncommon activity frame if the distance exceeds a certain limit of the constant threshold value. Next, we will go over the findings of the experiment.

Algorithm: Motion information retrieval algorithm for spotting anomalous human activities in video.

Inputs: ***Vi**, the surveillance video input*
***N**, No. of frames in video*
***W**, Width of the block*
***S**, Size of block for each frame*
***f**, the sequence of frame*
***Blk**, the individual block within the frame*
Output: ***Mof**, the motion feature vector*

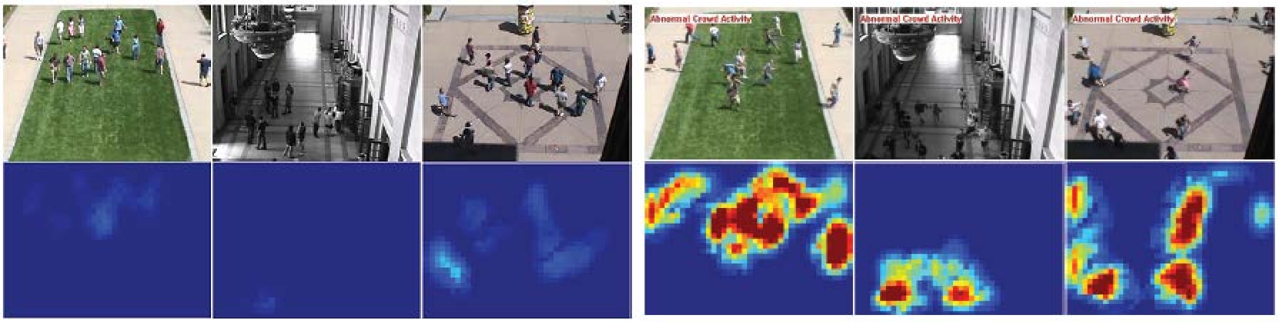
[Read sequence of frame from video]
***for f* = 1** to ***N* do**
  [Compute for each block Blk]
  ***for i* = 1** to ***S* do**
    [Compute optical flow at the pixel level]
    $(u, v) \leftarrow OpticalFlow\ (Blk_i)$
    [Compute each block threshold]
    $Th_d \leftarrow Blk_i\ x\ W$
      [Process neighbor frames in video sequence]
      for j* = 1** to ***K do
        [Compare the dissimilar frame in video sequence]
        If $i \neq j$ ***then***
          [Compute the neighbor block Euclidean distance]
          $EuD_{(i,j)} \leftarrow EuD(Blk_i, Blk_j)$
          [Compute distance matrix and compare with threshold]
          if $|\ EuD_{(i,j)} < Th_d|$ **then**
            Compute Angle Θ_{ij} *across adjacent block* Blk_i *and* Blk_j
            if $-\Theta Blk_i < \Theta_{ij} < \Theta Blk_j$ **then**
              $MoF \leftarrow MoF(\Theta Blk_i) + EuD_{(i,j)}/Blk_i$
            end if
          end if
        end if
    end for
  end for
end for

4. Experimental Results

In the experiment, datasets for proposed student behaviour, anomaly detection datasets from the University of Minnesota, and public databases were all used. Two different factors were taken into account: 1) In the proposed and University of Minnesota datasets for anomaly identification, identifying global unusual activity at the frame level; and 2) identifying local odd activity at the and pixel levels. In this case, the pixel level detection accuracy was assessed using the fact that the localized area's - which is regarded as exceptional - overlapping size is greater than 40% of the ground truth. Similarly to earlier studies [31, 32, 33], we used the parameters of the receiver operating characteristics (ROC) curve, the area under the ROC curve (AUC), and the equal error rate (EER).

4.1. University of Minnesota Pedestrian Dataset

The University of Minnesota dataset includes 11 video clips of crowded escape situations from three separate indoor and outdoor environments. There are 7,740 frames total, each measuring 320x240 pixels. A person walks around the scene at the beginning of each video segment. In this instance, walking was regarded as a typical activity (Fig.3a). Everyone in each clip makes an abrupt, quick-moving movement away from the scene (Fig.3b), which was considered unique in this collection. Walking along a walkway was expected to be a typical activity for this dataset. Ground truths at the pixel and frame levels are provided by the dataset, allowing for the localization of the unexpected activity-rich areas. 34 practice clips and 36 exam segments make up single pedestrian. It has a 238x158 frame size. There are 16 training video and 12 test clips in multiple pedestrian. The frame is 360x240 in size. The training videos only show regular pedestrians moving along a walkway. Unusual activities including bicycling, skating, and cart-driving are being done for the test clips.



(a) (b)

Fig. 3. Sample frames of University of Minnesota dataset and corresponding motion based Optical flow (a) For normal activity. (b) For unusual activity.

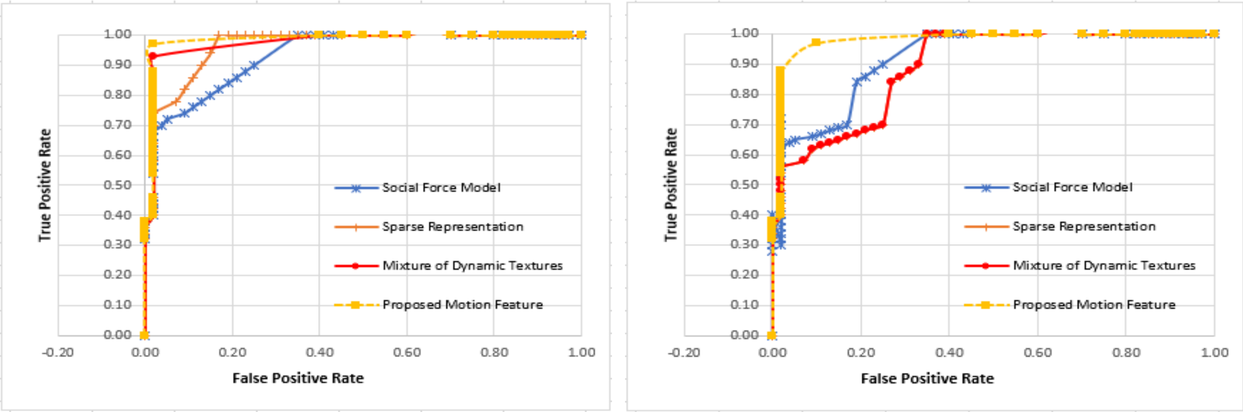
4.2. Proposed Pedestrian Dataset

The student's behavior on college property was observed using a top-notch DSLR camera from a different vantage point for the proposed dataset. We used a resolution of 3840x2160 and H.264 compression for shooting video at 30 frames per second. Around 100 sample videos are present in the database. Each example video lasts 20 to 30 minutes. In the dataset, there is a video series with an example frame shown in Fig.4. The camera tilts from 45 to 90 degrees. Over 90% of Yeshwantrao Chavan College of Engineering students in Nagpur are pedestrians. They are primarily of Indian origin, aged 22 to 27, and 65% of them are male. The proposed pedestrian database includes a variety of student actions in academic activities under various circumstances, such as students studying in practical laboratories, test hall scenarios, the classroom, a student cheating in the exam hall, a student removing the answer key outside the exam hall, a student stealing the mobile phone or other electronic devices like the mouse or keyboard, a student stealing the lab equipment, a student quarrel on the college grounds, etc. The proposed dataset includes expert human annotations for each video frame. For every video sequence, we produce a csv file with the exact same file naming convention as videofilename.csv. There are three phases to the labelling process: 1) Human detection; 2) tracking; 3) identifying unusual activities; and 4) soft biometric features Prior to manually validating and correcting the data, the Mask R-CNN [12] approach is used to produce a first estimate of each pedestrian's position in the scene. The preliminary tracking information was then provided using the deep sort method [14] and was once again rectified manually. We get a rectangular bounding box that represents the region of interest (ROI) for each pedestrian in each frame as a result of these first two stages. In the manual last stage of the annotation process, a human expert who is familiar with the college's students creates ID information and assigns soft labels to the samples. We also give a list of potential values for each label.

Fig. 4. Proposed pedestrian dataset. The first row shows the original sample images in the proposed pedestrian database. The second row shows the pedestrian annotation.

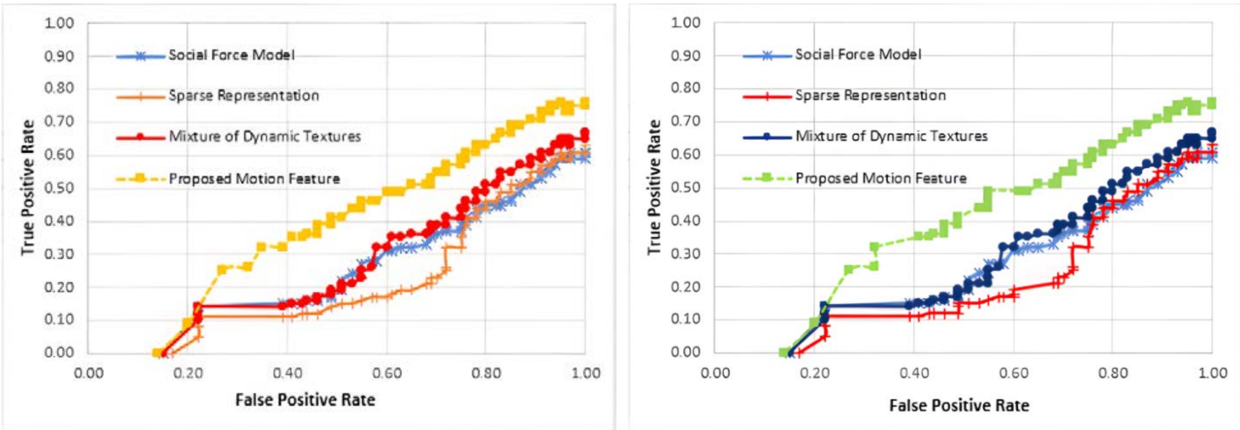
4.3. Performance Evaluation

Each video frame is split into eight 8x8-pixel blocks, and the threshold is set to the motion feature's maximum feature value in the training frames of the input video sequence. Receiving Operating Curves (ROC) for the proposed and current approaches are shown in Fig. 5 and are taken from [31, 32, 33]. We analyze the proposed method's ROC curve with the ROC curves of three cutting-edge approaches: sparse reconstruction [31], a social force model [32], and a mixture of dynamic textures [33].



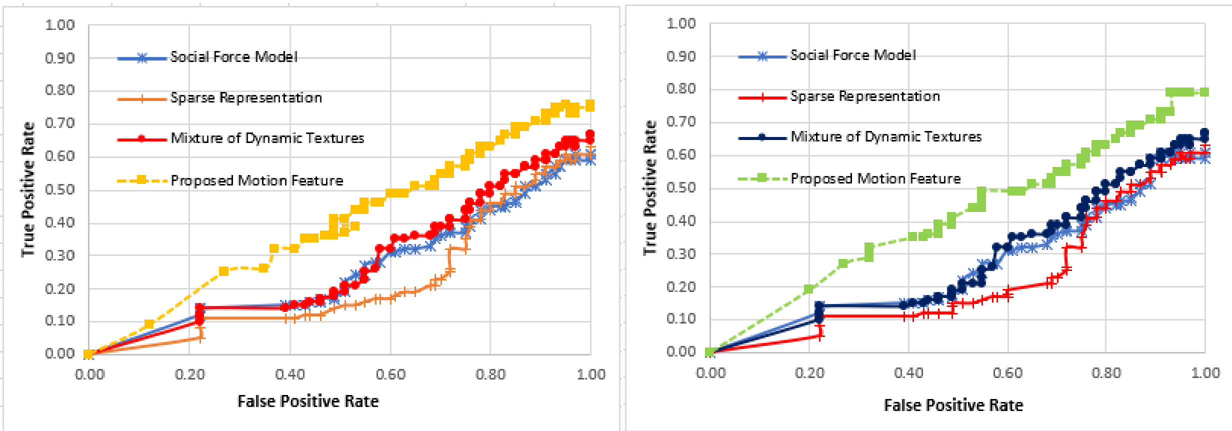
(a) (b)

Fig. 5. ROC curve for the comparative analysis (a) For USD dataset. (b) Proposed Dataset.



(a) (b)

Fig. 6. (a) Frame Level local anomalous activity detection (a) For USD dataset. (b) For proposed dataset.



(a) (b)

Fig. 7. (a) Pixel level local anomalous activity detection (a) For USD dataset. (b) For proposed dataset.

The proposed and current approaches given in [31, 32, 33] are represented by the Receiving Operating Curve (ROC) in Fig. 6 at the frame-level and pixel-level in Fig.7. In Tables 2 and 3, which provide quantitative comparisons, we also provided the ERRs and AUCs of the competing approaches, which were derived from the corresponding studies. On both datasets, with respective error rates of 16.1% and 18.1%, we found that the proposed technique performs comparably well compare to the conventional methods.

Table 2. Comparison of the proposed methodology with the state-of-the-art method on University of Minnesota anomaly detection dataset (ERR) frame level.

Technique or Methodology	Single Pedestrian	Multiple Pedestrian	Average
Dynamic Texture Mixture [32]	22.9%	22.9%	22.9%
Social Force Model [32]	36.5%	35.0%	35.7%
Sparce Matrix Representation [31]	35.6%	35.8%	35.7%
Motion-based method (Proposed)	21.1%	18.1%	16.1%

Table 3. Comparison of the proposed methodology with the state-of-the-art method on proposed dataset (AUC) at frame-level.

Technique or Methodology	Single Pedestrian	Multiple Pedestrian	Average
Dynamic Texture Mixture [32]	24.9%	23.9%	23.5%
Social Force Model [32]	37.5%	34.0%	36.7%
Sparce Matrix Representation [31]	32.6%	35.8%	33.7%
Motion-based method (Proposed)	22.1%	19.2%	18.1%

As demonstrated in Tables IV and V for both datasets, Area Under the Curve (AUC) for the current approach and the proposed method, respectively. The AUC for the proposed approach is 73.2% and 72.1%. It demonstrates that the suggested framework is more reliable and accurate when applied to the dataset and is, in comparison, superior to other methods of addressing anomalous behavior. Using a straightforward distance threshold-based classification algorithm, our method produced results on par with cutting-edge techniques.

Table 4. Comparison of the proposed methodology with the state-of-the-art method on University of Minnesota anomaly detection dataset (AUC).

Technique or Methodology	Single Pedestrian	Multiple Pedestrian	Average
Dynamic Texture Mixture [32]	59.3%	56.8%	58.0%
Social Force Model [32]	40.9%	27.6%	34.2%
Sparce Matrix Representation [31]	32.6%	22.4%	27.5%
Motion-based method (Proposed)	64.9%	81.5%	73.2%

Table 5. Comparison of the proposed methodology with the state-of-the-art method on proposed dataset (AUC).

Technique or Methodology	Single Pedestrian	Multiple Pedestrian	Average
Dynamic Texture Mixture [32]	35.6%	35.8%	35.7%
Social Force Model [32]	50.8%	63.4%	57.1%
Sparce Matrix Representation [31]	74.5%	70.1%	72.3%
Motion-based method (Proposed)	63.4%	80.2%	72.1%

5. Conclusion

In order to identify students engaging in aberrant behavior in a classroom setting, this paper offers a practical and original methodology. We can characterize pedestrian activity in a frame as normal or abnormal, and we can also pinpoint specific portions of abnormal activity as local or global regions, according to the spatial and temporal characteristics of motion-based features. We conducted tests using proposed student behavior datasets and anomaly detection datasets from the University of Minnesota. In comparison to state-of-the-art techniques, the experimental results demonstrate that the suggested framework lowers the false acceptance rate to 3% and raises the True positive rate to 5%. The primary objective of this effort is, however, to effectively identify aberrant behavior in a classroom setting. The similar approach can be utilized in the future for a variety of scenarios involving student conduct, including situations when students are caught cheating on exams or getting into arguments on campus. Again, if the proposed approach is improved by the inclusion of further features like scale, rotation, and lighting invariant features, scale, rotation, and illumination changes can also be addressed.

References

[1] Wanli Ouyang, Xingyu Zeng, Xiaogang Wang, Single-Pedestrian Detection Aided by Two-Pedestrian Detection, *TPAMI*, vol. 37, no. 9, pp. 1875-1889, September 2015.

[2] Weihang Kong, He Li, Guanglong Xing, Fengda Zhao, "An Automatic Scale-Adaptive Approach With Attention Mechanism-Based Crowd Spatial Information for Crowd Counting," in *IEEE Access*, vol. 7, pp. 66215-66225, 2019.

[3] Khosro Rezaee, Mohammad R. Khosravi, Maryam Saberi Anari, "Deep-Transfer-Learning-Based Abnormal Behavior Recognition Using Internet of Drones for Crowded Scenes," in *IEEE Internet of Things Magazine*, vol. 5, no. 2, pp. 41-44, June 2022.

[4] Lei Cao and Lin Wang, "Analysis and intervention in pedestrian evacuation via a modified social force model," *The 27th Chinese Control and Decision Conference (2015 CCDC)*, 2015, pp. 5924-5929, doi: 10.1109/CCDC.2015.7161870.

[5] Fanchao Xu, Yunbo Rao, Qifei Wang, "An unsupervised abnormal crowd behavior detection algorithm," *2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, 2017, pp. 219-223, doi: 10.1109/SPAC.2017.8304279.

[6] Fabian Flohr, Madalin Dumitru-Guzu, Julian F. P. Kooij, Dariu M. Gavrila, “A Probabilistic Framework for Joint Pedestrian Head and Body Orientation Estimation,” in *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1872-1882, Aug. 2015, doi: 10.1109/TITS.2014.2379441.

[7] Amir Rasouli, Iuliia Kotseruba, John K. Tsotsos, "Understanding Pedestrian Behavior in Complex Traffic Scenes," in *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 1, pp. 61-70, March 2018, doi: 10.1109/TIV.2017.2788193.

[8] Liang-qun Li, Chun-lan Li, Jun-bin Liu, Wei-xin Xie, "A novel occlusion handling method based on particle filtering for visual tracking," 2016 *IEEE 13th International Conference on Signal Processing (ICSP)*, 2016, pp. 934-937, doi: 10.1109/ICSP.2016.7877967.

[9] Bing Wang, Kap Luk Chan, Gang Wang, Haijian Zhang, “Pedestrian detection in highly crowded scenes using “online” dictionary learning for occlusion handling,” *2014 IEEE International Conference on Image Processing (ICIP)*, 2014, pp. 2418-2422, doi: 10.1109/ICIP.2014.7025489.

[10] Ho-chul Shin, "Abnormal Detection based on User Feedback for Abstracted Pedestrian Video," *2019 International Conference on Information and Communication Technology Convergence (ICTC),* 2019, pp. 1036-1038, doi: 10.1109/ICTC46691.2019.8939600.

[11] Md Maruf Hossain Shuvo, Nafis Ahmed, Koundinya Nouduri, Kannappan Palaniappan, "A Hybrid Approach for Human Activity Recognition with Support Vector Machine and 1D Convolutional Neural Network," *2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 2020, pp. 1-5, doi: 10.1109/AIPR50011.2020.9425332.

[12] Arundhati Chattoadhyay, Aruna Tripathy, “Anomalous Crowd Behavior Detection Using Raspberri PI,” *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA),* 2020, pp. 462-466, doi: 10.1109/ICCCA49541.2020.9250824.

[13] Kazunari Takasaki, Ryoichi Kida, Nozomu Togawa, “An Anomalous Behavior Detection Method Based on Power Analysis Utilizing Steady State Power Waveform Predicted by LSTM,” *2021 IEEE 27th International Symposium on On-Line Testing and Robust System Design (IOLTS)*, 2021, pp. 1-7, doi: 10.1109/IOLTS52814.2021.9486706.

[14] Feng Zheng, Quanyun Liu, “Anomalous Telecom Customer Behavior Detection and Clustering Analysis Based on ISP’s Operating Data,” in *IEEE Access*, vol. 8, pp. 42734-42748, 2020, doi: 10.1109/ACCESS.2020.2976898.

[15] Shuoyan LIU, Hao XUE, Chunjie XU, Kai FANG, "Abnormal behavior detection based on the motion-changed rules," *2020 15th IEEE International Conference on Signal Processing (ICSP)*, 2020, pp. 146-149, doi: 10.1109/ICSP48669.2020.9321012.

[16] Wei Fang, Lin Wang, Peiming Ren, "Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments," in *IEEE Access*, vol. 8, pp. 1935-1944, 2020, doi: 10.1109/ACCESS.2019.2961959.

[17] Zhihong Sun, Jun Chen, Liang Chao, Weijian Ruan, Mithun Mukherjee, “A Survey of Multiple Pedestrian Tracking Based on Tracking-by-Detection Framework,” in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 5, pp. 1819-1833, May 2021, doi: 10.1109/TCSVT.2020.3009717.

[18] Yang Li, Jinshen Wang, Xiang Liu, Ning Xian, Changsheng Xie, "DIM Moving Target Detection using Spatio-Temporal Anomaly Detection for Hyperspectral Image Sequences," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium,* 2018, pp. 7086-7089, doi: 10.1109/IGARSS.2018.8517601.

[19] Yunpeng Jiang, Ningning Le, Yufeng Zhang, Yinger Zheng, Yang Jiao, "Research on the Flight Anomaly Detection During Take-off Phase Based on FOQA Data," *2019 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS),* 2019, pp. 756-760, doi: 10.1109/SAFEPROCESS45799.2019.9213422.

[20] Chaoqin Huang, Qinwei Xu, Yanfeng Wang, Yu Wang, Ya Zhang, "Self-Supervised Masking for Unsupervised Anomaly Detection and Localization," in *IEEE Transactions on Multimedia*, doi: 10.1109/TMM.2022.3175611.

[21] Gal Mishne, Israel Cohen, "Multiscale Anomaly Detection Using Diffusion Maps," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 1, pp. 111-123, Feb. 2013, doi: 10.1109/JSTSP.2012.2232279.

[22] Kursat Rasim Mestav, Xinyi Wang, Lang Tong, "A Deep Learning Approach to Anomaly Sequence Detection for High-Resolution Monitoring of Power Systems," in *IEEE Transactions on Power Systems*, doi: 10.1109/TPWRS.2022.3168529.

[23] Huawen Liu, Xiaodan Xu, Enhui Li, Shichao Zhang, Xuelong Li, "Anomaly Detection With Representative Neighbors," in *IEEE Transactions on Neural Networks and Learning Systems*, doi: 10.1109/TNNLS.2021.3109898.

[24] Chunde Yang, Tongyao Xu, Minghong Lv, Yi Yu, Xiaoming Jiang, "Pedestrian Angle Recognition Based on JDE Multi-object Tracking Algorithm," *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 2022, pp. 647-651, doi: 10.1109/ICSP54964.2022.9778590.

[25] Chuangxin He, Xu Zhang, Zhonghua Miao, Teng Sun, "Intelligent vehicle pedestrian tracking based on YOLOv3 and DASiamRPN," *2021 40th Chinese Control Conference (CCC)*, 2021, pp. 4181-4186, doi: 10.23919/CCC52363.2021.9549997.

[26] Zhihong Sun, Jun Chen, Liang Chao, Weijian Ruan, Mithun Mukherjee, "A Survey of Multiple Pedestrian Tracking Based on Tracking-by-Detection Framework," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 5, pp. 1819-1833, May 2021, doi: 10.1109/TCSVT.2020.3009717.

[27] Kai Chen, Xiao Song, Xiang Zhai, Baochang Zhang, Baocun Hou, Yi Wang, “An Integrated Deep Learning Framework for Occluded Pedestrian Tracking," in *IEEE Access*, vol. 7, pp. 26060-26072, 2019, doi: 10.1109/ACCESS.2019.2900296.

[28] Shan Gao, Zhenjun Han, Ce Li, Qixiang Ye, Jianbin Jiao, “Real-Time Multipedestrian Tracking in Traffic Scenes via an RGB-D-Based Layered Graph Model,” in *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2814-2825, Oct. 2015, doi: 10.1109/TITS.2015.2423709.

[29] Ying Liu, Weiwei Zhang, Guohua Cui, Yong Zhang, “Multipedestrian Online Tracking Based on Social Force-Predicted Deformable Key-Points Mapping via Compressive Sensing,” in *IEEE Systems Journal*, vol. 15, no. 2, pp. 1905-1916, June 2021, doi: 10.1109/JSYST.2020.3003024.

[30] Qinglin Tian, Zoran Salcic, Kevin I-Kai Wang, Yun Pan, “A Multi-Mode Dead Reckoning System for Pedestrian Tracking Using Smartphones,” in *IEEE Sensors Journal*, vol. 16, no. 7, pp. 2079-2093, April1, 2016, doi: 10.1109/JSEN.2015.2510364.

[31] Zhengqiang Jiang, Du Q. Huynh, “Multiple Pedestrian Tracking From Monocular Videos in an Interacting Multiple Model Framework,” in *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1361-1375, March 2018, doi: 10.1109/TIP.2017.2779856.

[32] M. Zaharin, N. Ibrahim, T. Tuan Dir,. “Comparison of human detection using background subtraction and frame difference”. *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 1, pp. 345-353, 2020 , doi: 10.11591/eei.v9i1.1458.

[33] M. Gheisari, Z. Safari, M. Almasi, A. Najafabadi, A. Sridharan, R. G K, Y. Liu, A. Abbasi,. "A novel enhanced algorithm for efficient human tracking". *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 11, no. 1, pp. 1-7, 2022 , doi: 10.11591/ijict.v11i1.pp1-7.

[34] S. V. Aruna Kumar, Ehsan Yaghoubi, Abhijit Das, B. S. Harish, Hugo Proença, "The P-DESTRE: A Fully Annotated Dataset for Pedestrian Detection, Tracking, and Short/Long-Term Re-Identification From Aerial Devices," in *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1696-1708, 2021, doi: 10.1109/TIFS.2020.3040881.

[35] L. Knoedler, C. Salmi, H. Zhu, B. Brito and J. Alonso-Mora, "Improving Pedestrian Prediction Models With Self-Supervised Continual Learning," in *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4781-4788, April 2022, doi: 10.1109/LRA.2022.3148475.

[36] G. Shen, L. Zhu, J. Lou, S. Shen, Z. Liu, L. Tang, "Infrared Multi-Pedestrian Tracking in Vertical View via Siamese Convolution Network," in *IEEE Access*, vol. 7, pp. 42718-42725, 2019, doi: 10.1109/ACCESS.2019.2892469.

[37] W. Ge, R. Collins, R. Ruback, "Vision-Based Analysis of Small Groups in Pedestrian Crowds," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 5, pp. 1003-1016, May 2012, doi: 10.1109/TPAMI.2011.176.

[38] M. Rai, R. Yadav, A. Husain, T. Maity, D. Yadav,"Extraction of Facial Features for Detection of Human Emotions under Noisy Condition", *International Journal of Engineering and Manufacturing*, Vol.8, No.5, pp.49-62, 2018.

[39] F. Chamasemani, L. Affendey, "Systematic Review and Classification on Video Surveillance Systems", *I.J. Information Technology and Computer Science*, Vol. 2013, No.7, pp.87-102, June 2013.

[40] A. Dey, F. Mohammad, S. Ahmed, R. Sharif, A. Saif, "Anomaly Detection in Crowded Scene by Pedestrians Behaviour Extraction using Long Short Term Method: A Comprehensive Study", *International Journal of Education and Management Engineering*, Vol.9, No.1, pp.51-63, 2019.

[41] Y. Imamverdiyev, L. Sukhostat, "An Anomaly Detection Based on Optimization", *International Journal of Intelligent Systems and Applications*, Vol.9, No.12, pp.87-96 2017.

[42] H. Salami, R. Ibrahim, M. Yahaya, "Detecting Anomalies in Students' Results Using Decision Trees", *International Journal of Modern Education and Computer Science*, Vol.8, No.7, pp.31-40, 2016.

Authors' Profiles

Kamal O. Hajari is pursuing a Ph.D. degree from the IT Department, YCCE. He completed his bachelor and master degree from YCCE College. It is an Autonomous institution affiliated to Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur. His research interests of areas are video surveillance, video processing and analysis, pattern recognition, machine learning, neural network and fuzzy logic, and computer vision. He is a life time member of Institute of Engineers (IE), ACM, and ISTE.

Ujwalla H. Gawande is an Associate Professor and Dean R & D, IT Department, YCCE. She got her B.E. from KITS, Ramtek and M.Tech. Degree from Raisoni college of Engineering, RTMNU University, Nagpur in 2001 and 2007, and a Ph.D. degree from SVNIT, Surat in 2014. Her research interests of areas are video surveillance, video processing and analysis, pattern recognition, machine learning, neural network and fuzzy logic, and computer vision. She is a reviewer of peer review journal such as AI, IEEE Transaction of Multimedia, and IEEE Access. She is having life time membership of IE, ACM, and ISTE.

Yogesh Golhar is an Assistant Professor, CSE Department, St. Vincent Palloti College of Engineering and Technology, Nagpur. He got his BE and M.Tech. Degree from Raisoni College in 2008 and 2012. His research area includes AI, Big Data Analytics, ML, CN, and CV. He is having membership of IE, ACM, and ISTE.

How to cite this paper: Kamal Omprakash Hajari, Ujwalla Haridas Gawande, Yogesh Golhar, "Motion Pattern Based Anomalous Pedestrian Activity Detection", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.14, No.6, pp. 15-25, 2022. DOI:10.5815/ijigsp.2022.06.02