Human Body Fall Recognition System

Jannatul Mourey,[1] Ava Sehat Niaki[1], Rupali Gupta[1], Priyanka Kaplish[1]

Department of Computing Science, University of Alberta, Canada
mourey@ualberta.com,
asehatni@ualberta.ca,rupali1@ualberta.ca,kaplish@ualberta.ca

Abstract. Falling is one of the major risks in elderly people, kids or people with disabilities. The situation gets bitter when the victim undergoes serious injuries and is not able to get the help on time. In this paper, we are proposing a method to detect human body fall in real-time environment. The proposed detection algorithm consists of three stages: Video analysis, Human body recognition and trigger alert. We have implemented the method using OpenCV and Python libraries. This approach has been tested in the series of environments and the experimental results mentioned in this report shows impact-full outcomes leveraging different accuracy methods.

Keywords: Fall Detection, Video analysis, Human body recognition, Background Subtraction, OpenCV.

1 Introduction

Activity recognition is an active research topic in computer-vision. Visual surveillance is majorly used to monitor the movements of people. Automated systems can be developed to store the information and analyze it later. Those can be used to capture the human fall and report it as soon as possible. If the fall is not met with immediate help, it may cause some serious injuries due to late medical salvages or if the person remains immobilized or unconscious for a longer duration. We need to develop new technologies to ensure the safety of people suffering from fall-related injuries. Surveillance technologies can help detecting the fall and help mitigate the injuries. Videos captured from surveillance cameras can be analyzed detect the fall and get help on time. Video based fall detection goes through three stages: Video acquisition, video analysis and notification. The proposed system gives human activity detection from the streamed video which will be analyzed to distinguish the moving object(human) from the background. Consequently, the system will extract the important information of the moving object, and check if it is classified as falling down event or a normal motion. This system is implemented for real-time environment. There are various sensor based techniques, which use sensors like pressure, acceleration and sound that have been used over the period of time to detect a fall; the major advantage of vision-based systems over these is that a person does not have to wear the sensor all the time. In case the person forgets to wear the sensor, a camera can always monitor the person. Another advantage of a vision-based system is that it provides considerable data to be processed. The detection algorithm can analyze- using a combination of the images from the actual fall- that how serious the fall was. Through cameras, we can calculate how much time the person took to get up or if the person was able to get up at all. But if a sensor fails to capture the fall, it has no second chance at detecting the fall. Apart from detecting the fall, a single camera can detect other abnormal activities too.

2 Related Work

Majority of the fall detection systems can be classified into three categories: Wearable sensor-based, ambient sensor-based and computer vision methods. Wearable sensor based method uses an accelerometer sensors which is attached to subject's body to capture the motion. Ambient sensor-based method uses external sensors including pressure sensors, acoustic sensors and so on [1]. In the Computer-Vision based method, images and videos are analyzed to detect.[2][3][4][5][6][7]

- Sensors based approach:External sensors are deployed on the subject of interest(SOI). Accelerometer is the most widely used wearable sensor for fall detection. The wearable sensor approach infers the detection based on embedded sensors to sense the movement and position of the object.
- Moment functions and depth image in fall detection: The most common moment functions used in computer vision applications are shape analysis, shape deformation description, postures estimation, and so on. These are used for ellipse fitting to represent shape deformation, and then silhouette motion is modelled for fall detection. The features of extracted human silhouettes are described by the shape and position of the silhouette. Support Vector Machine (OCSVM) Method and Multi-class SVM Classification can be used to classify the normal daily postures and other human activities or falls. Moment functions can be used to describe the human shape which is approximated by an ellipse which can be best used for human body recognition. Ellipse approximation and bounding box- both are adapted in our work for the best results under different environments.
- Classification of fall detection methods:Classification of the fall is done based on the fall definition. It is used to distinguish between the fall and other human activities like bending, sitting, tripping etc. The collected data is distinguished based on the feature extraction and selection. Decision Trees (DTs) are one of the oldest algorithms used for pattern classification. Artificial Neural Networks (ANNs) are also used to classify falls from daily

activities. Support Vector Machine(SVM) can also be used to classify different postures to identify the fall. A Bayesian Belief Network (BBS) can be used to model the causality of the slip or fall events with other events. Fall detection can be carried out in two stages: fall definition and detection. In Artificial Intelligence approach, model is trained to learn the patterns and distinguish between the patterns. However, a lot of training is required for the training phase. To avoid this, threshold methods can be used instead of learning methods. Threshold-based method includes detecting falls by measuring the distance and angle between centroids of human body and the floor, and setting it as a threshold. Some methods use head-shoulder gap and head-ground gap for thresholding the values. If the orientation is smaller than than the threshold, then the fall is detected.

3 Implementation Details

The proposed approach in this project takes into account various measures derived by the environment and human body activities. Based on environmental factors of the test data-sets, we have pre-processed the data-set to get the better understanding of human position in different surroundings. Following that, we have studied behavioral elements of human body to predict a fall based on human body response to a sudden unintentional movement. To develop and test this approach, we have commenced our research by implementing the Le2i data-set [8].

3.1 Data-set

Le2I data-set is created specifically for assessing human fall in different surroundings. This data-set includes various scenarios of human activities which are not limited to a simple fall; these activities include and are not limited to sitting, resting and bending. In addition to assorted activities, the mentioned data-set incorporates various environments with myriad challenges. In this project we have tested our method on two environments-office and home. It is worth mentioning that these particular environments cover distinct illumination and motion interference trials in series of more that 30 videos per environment.

The proposed approach is developed for accurate classification of activities in the Le2i videos and eventually identifying the fall. This method consists of five stages,: Video analysis, Human body approximation, Fall detection and notification. Following the proposed architecture (Figure 1), each stage is elaborated further in the paper.

- Video Analysis : The first step of the proposed approach involves segmentation of the video into frames. An illumination rate for the initial captured frame is calculated based on average frame pixel value of the gray-scale conversion of the video which is used to threshold the illumination of the video. To ensure the best results for background subtraction, we are using two methods: GMG and MOG2, for background subtraction.

  - The GMG algorithm uses first few frames for background modelling. It employs probabilistic foreground segmentation algorithm that identifies possible foreground objects using Bayesian inference.This algorithm combines statistical background image estimation and per-pixel Bayesian segmentation.
  - MOG2 is a Gaussian Mixture-based Background/Foreground Segmentation Algorithm. The weights of the mixture represent the time proportions that those colours stay in the scene and selects the appropriate number of gaussian distribution for each pixel.

From the mentioned methods, an optimal method is selected based on the motion and illumination rate calculated earlier(MOG2 and GMG background subtractors are triggered based on the motion estimation and illumination threshold). To reduce unwanted noise in each frame, a series of filters such as blurring and averaging filters are applied. By the end of this step, the algorithm creates a binary mask frame of the moving object in the video.

– Human body approximation:. After pre-processing the frame and careful selection of Region of Interest(ROI), the human body is approximated with a bounding box and an ellipse. Human body measurements are then thresholded to differentiate between human body and other objects.
– Fall Definition: Fall is defined based on the measured parameters of bounding boxes and ellipse. When the angle between major axis and floor and the size of ellipse, or the ratio of width to height of bounding box goes beyond the threshold it is considered as a fall. In figure 3, initiation and period of the fall is represented for three sample videos. These graphs are a parallel representation of thee measured values and how they change when a fall happens. By studying the changes in boundary box ratio(w/h), ellipse orientation(theta) and size of the major axis of ellipse(f), we set thresholds for defining the fall. Fall can be detected in three scenarios:
  - Change of w/h: In this case as the size of human body approximation changes from vertical to horizontal(w/h¿1), fall is detected. This method is most viable when a person falls to the sides. The ellipse size in this method is a discriminating approach to identify the fall.
  - Change of theta and f: As theta gets closer to pi or 0 degrees, the change in the angle is used to detect the fall. It can be used when human falls towards the camera.
  - In some scenarios, both the methods can be used to define the fall.
– Fall detection: Leveraging the mentioned measures, the proposed algorithm is set to record the frame details, absolute values of theta, f and w/h. The contained data is consequently used for measuring the period and intensity

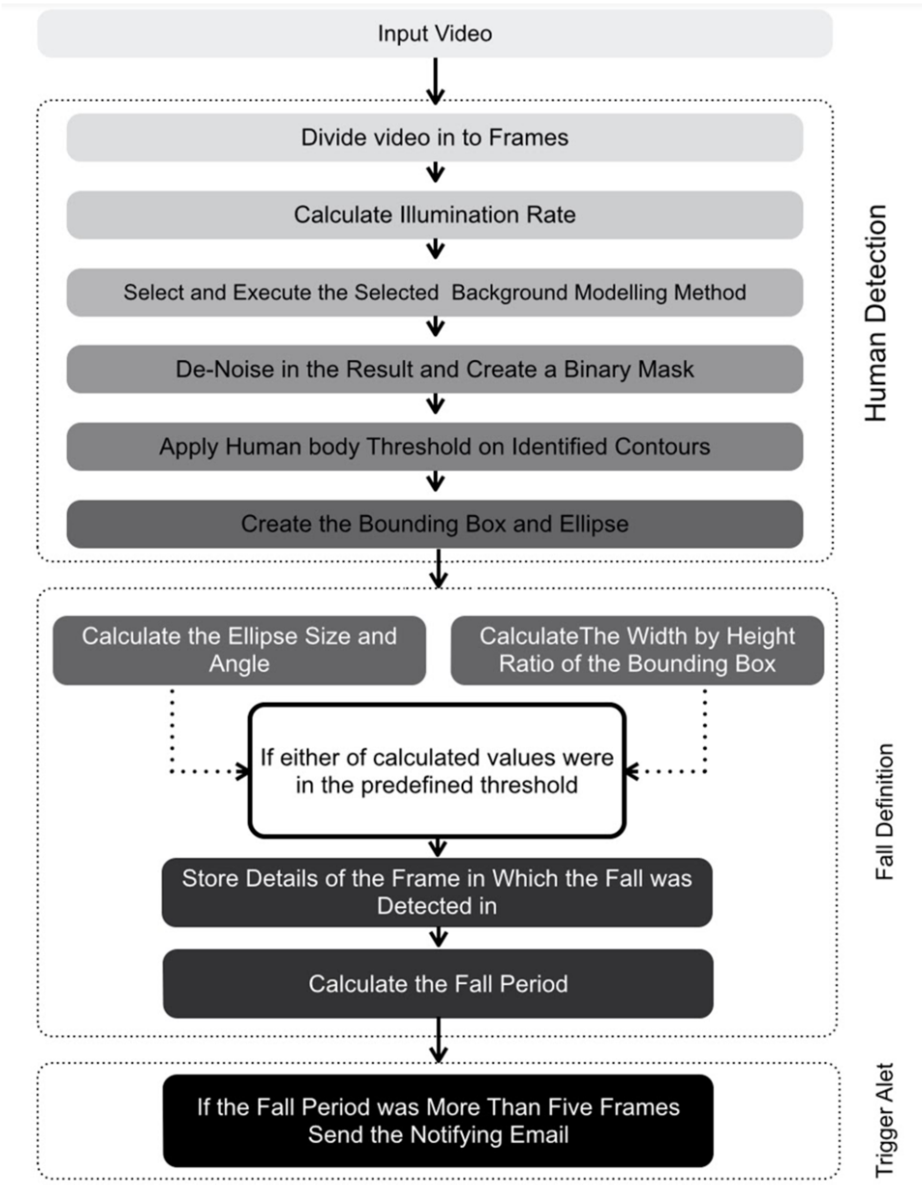

Fig. 1. System Architecture

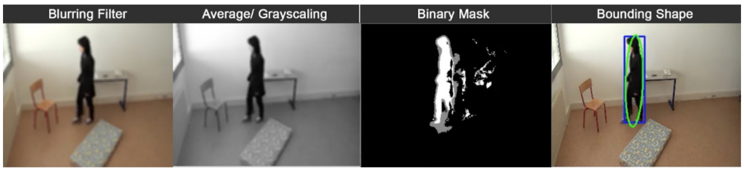

Fig. 2. Frame Processing Step by Step

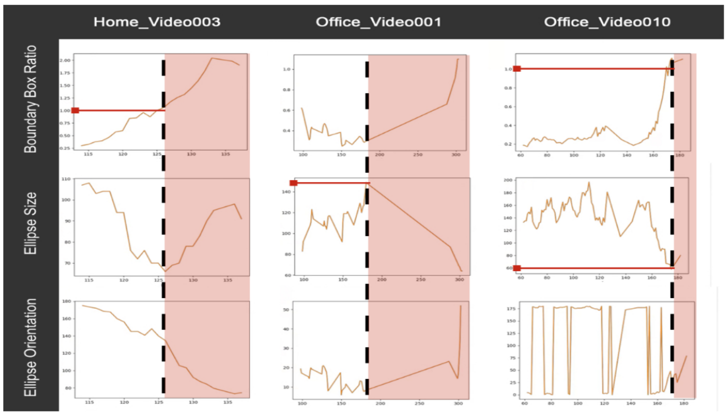

Fig. 3. Fall Detection By Value Analysis

of the fall. For the purpose of visual readability, the shape of the human body approximation is assigned a different colour as the fall is detected(Figure 4).

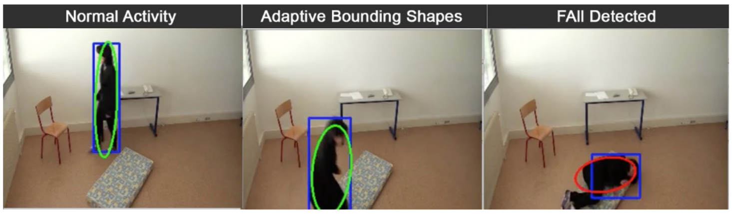

Fig. 4. Detected Fall

- Notification: After the fall is been defined and detected, we capture the frame details in which the fall happened and register the fall period. Finally, a notifications is sent if the fall period is more than 5 frames.

3.2 Results

The accuracy of the proposed system is measured based on two methods:

- Method 1: This method is designed based on categorizing the test outcomes into four types. The precision of the method is measured using the mAP (mean Average Precision) for Object Detection technique introduced by Perry, Kent Berry. Precision measures how accurate are the predictions, it calculates the percentage of positive predictions.

tableAccuracy Method 1

Accuracy Calculation Method 1				
	True- Positive	False- Negative	False- Positive	True- Negative
Action	Fall and detected	fall Not Detected	No Fall but Detected	Fall Not Detected

- Method 2:This method relies on the potential of the selected data-set. An annotation file is provided with the data-set which contains the information of the fall detected. By comparing our results with the details provided in the file, we calculate the accuracy of our approach.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

Based on the above mentioned methods, our approach was able to distinguish between a fall and fall-resembling action in the tested videos successfully. The current accuracy rate of the proposed approach is 85% (accuracy method 1) and 86.6% (accuracy method 2).

tableAccuracy Results

Approach Accuracy						
	Average Accuracy (Method 1)	Average Accuracy (Method 2)	Office (Method 1)	Office (Method 2)	Home (Method 1)	Home (Method 2)
Accuracy	83.33%	86.67%	83.33%	86.6%	86.67%	N/A

3.3 Limitations and Challenges

Even though human fall detection is an widely discussed and developed topic, proposed approached faces various challenges. These challenges are categorized in three areas: a)true detection of human, b)differentiating between human activities in order to define a fall and c)creating an up-to-date informative notification system.

- Identification of Human Body :To insure an accurate fall detection, the first step is to identify proper segmented foreground/background. Challenges in identifying the accurate foreground/background can be categorized in four areas:
  - Changes in Illumination: Illumination rates are the key factor in concise contouring of the human body; in a very dark or light environment defining the measurements of human silhouette can be challenging which can lead to misidentification. In order to achieve a working identification algorithm for environments with variable illumination, our approach follows an elimination path to select the best background subtractor amongst the above mentioned models for a specific environment.
  - Other Moving Objects: in Frames Including Shadows: Processing a real-time scenario is more challenging than a simulated environment such as Le2i. Multiple moving object such as pets, furniture or shadows modulate the accuracy of the system. The accuracy of the method can be compromised if more False-Positives(detecting a false fall)are detected than the True-Positives. Our approach addresses this problem by pre-processing each frame of the input video and applying denoising filters,

in addition to converting it to grayscale, to simplify the frame before applying contours.
  - Distance From Camera: In the proposed approach, we identify a human body based on its proportion and bounding area. The predefined threshold, however, could not be as accurate in certain videos with higher depth of view(the size of a human in proportion to frame changes significantly). One solution to overcome this issue is to include other background subtraction methods in the preprocessing stage which address specific human features for recognition. In the scope of this project, we have implemented and tested HOG(feature-based object identification model) in the development phase. However this model is limited to certain human shapes and angles.
  - Occlusion: Another challenge is when human body to be detected is covered by other objects. The measures taken to in previous challenge can be used to detect a human in this feature limitation.
- Challenges in Fall Detection Limitation in defining a fall can rise either from the activity casting conditions or the activity itself. Position of a human body with respect to camera while falling and activities which result in sudden and intensive change in human proportion, are a few examples which limit capabilities of a detection system.
  - Other Human Activities Resembling Fall: Activities such as sitting, lying down and bending to pick up objects result in many challenges as the human body orientation and size changes significantly. The proposed method for fall detection is based on combination of the mentioned measures, therefore optimizing the process of differentiating between these activities.
  - Challenging Fall Angles : When a person falls down at certain angles with respect to the camera, the change in the orientation of body is not often detected by many fall detection systems. Even though the proposed model takes account of the changes in the ratio of human body, using a 3D images can provide better grounds for assessing body position in a real world [9]; we can take this research forward by investigating 3D images to optimize the current model.

Notification Following an accurate fall detection a responsive and operational notification system is required to ensure safety of the targets. Limitation of such system can be best as follows:
  - False Notification: Even by embedding various measures and methods for detecting the fall and defining a fall, false notification can get through and result in false alarm. In order to avoid such situation by measuring the period of fall, we have added a third measure-time, for defining a fall more precisely and avoid false notifications.
  - Late Notification: Late notifications can also be hinder the accuracy of the model. In many cases, the earlier the person who has undergone a

serious fall is recovered, higher the chances of recovering. Sending an alert email is one example of how the proposed fall detection system can communicate a fall. However, this means of notification can be replaced by a more convenient system according to the target environment; for instance in a hospital or a caring facility, employing a beeper would replace the current email system.

4 Conclusion

To conclude, this system has been developed while considering major factors affecting human body fall. We have used Gaussian mixture model to remove any unwanted noise from the environment. We have implemented two background subtraction method (MOG2 and GMG) to extract the foreground and background frame. After receiving the processed frame, the features of the human body fall are detected using the physics of bounding box and ellipse surrounding the target object. The purpose of using multiple geometrical figures is to get better accuracy without neglecting any external factors. We have also categorized the fall on the basis of its severity. The accuracy of this system is being calculated based on two different strategies in which it has successfully achieved 83.33 percentage (Method1) and 86.67 percentage (Method 2) accuracy in its result. All the external factors present in the data-set were considered during this process. The last phase of this system is the notification system in which a notification is triggered to the concerned person (guardian/caretaker) whenever the fall is detected. This application is very useful to develop industrial application for health care, construction sites, smart homes and so on.

References

1. Yang, L., R.Y..Z.W.: 3d depth image analysis for indoor fall detection of elderly people. **2**(1) (2016) 24–34
2. Alaliyat, S.: Video - based fall detection in elderly's houses. (2008)
3. Gaikwad, K., D. Patil, M.: Human activity detection and recognition algorithm from video surveillances. **2**(5) (2013)
4. Salmi, K.: Improving safety for home care patients with a low-cost computer vision solution. (2016)
5. SCHRADER, N.: Detecting falls and poses in image silhouettes. (2011)
6. Skolan i J. Šarlija, V., I.K.I.M.D.T.C.P., Krekovic, M.: A method for real-time detection of human fall from video. (2012)
7. Van Tuan, P., Le Uyen Thuc, H.: An effective video based system for human fall detection. **3**(8) (2014)
8. Le2i.cnrs.fr.: Fall detection dataset - le2i - laboratoire electronique, informatique et image. (2018)
9. Zhang, W., R.Y., Yang, L.: 3d depth image analysis for indoor fall detection of elderly people. **2**(1) (2006)