A Framework for Fall Detection Using Audio and Video Features

Medha Bisht
Department of Computer Science and Engineering,
Graphic Era (Deemed to be University)
Dehradun, India
medhabisht525@gmail.com

Aryamann Singh
Department of Computer Science and Engineering,
Graphic Era (Deemed to be University)
Dehradun, India
aryamann.singh21@gmail.com

Divyam Kholia
Department of Computer Science and Engineering,
Graphic Era (Deemed to be University)
Dehradun, India
kholiadivyam21@gmail.com

Yogesh Thakur
Department of Computer Science and Engineering,
Graphic Era (Deemed to be University)
Dehradun, India
11yogkur11@gmail.com

Durgaprasad Gangodkar
Department of Computer Science and Engineering,
Graphic Era (Deemed to be University)
Dehradun, India
dr.gangodkar@geu.ac.in

***Abstract*—Falls can lead to severe injuries, diminished quality of life, and, in unfortunate cases, even fatal consequences. As the elderly population grows worldwide, there is a demand for better surveillance systems, specifically fall detection systems to tackle this issue. The current methods commonly used to detect falls are portable sensors which need to be worn or embedded on various parts of the human body, in order for them to be able to detect falling events experienced by the person. This paper introduces an efficient fall detection system utilizing OpenCV, a powerful computer vision library. The system analyzes real-time video streams to detect falls among individuals, focusing on non-intrusiveness, real-time capability, and adaptability. Leveraging OpenCV's algorithms, the system distinguishes fall events from normal activities, demonstrating robustness against lighting variations and background clutter. Extensive testing validates its high accuracy and low false alarm rate, making it a promising tool for enhancing safety in environments with vulnerable populations.**

***Index Terms*—Fall detection System, Sensors, MediaPipe**

I. INTRODUCTION

As people get older, falling down becomes a big problem, especially for those 79 and older. Every year, about 684,000 people die from falls, according to the World Health Organization. To tackle this issue, we're working on a new way to quickly spot falls using both sound and video. Most fall detectors need you to wear something, but our idea uses cameras and sound instead. Cameras in good spots watch for falls, and we listen for sounds that show someone needs help. This makes our system not only smart but also easy to use without bothering people. Adding sound helps us understand better if someone is in trouble. Our paper dives into how we teach computers to understand what they see and hear in real-time and figure out how to handle tricky situations like low light or things blocking the view to make sure our system works well. But it's not just about tech stuff. We care about privacy and doing things the right way. We talk in our paper about being responsible with our system and making sure people's privacy is always respected. In a nutshell, our paper talks about a cool way to use cameras and sound to quickly notice when someone falls. We keep it simple and smart, aiming to make life better for older folks by getting them help fast.

II. LITERATURE SURVEY

This literature survey [10] discusses how Fall detection systems can be divided into 2 types: environmentally smart systems and wearable devices.

Early fall detection approaches from 2013 to 2017 focused on wearable sensors. Wearable Sensors utilize sensors placed on a person's body, such as accelerometers or gyroscopes, commonly found in fitness wearables or mobile phones. Accelerometers and gyroscopes integrated into wearable devices were widely explored to detect changes in movement patterns indicative of falls. However, user acceptance and comfort were major concerns [6,7].

On the other hand environmental systems employ external sensors like cameras, floor sensors, infrared sensors, microphones, or pressure sensors.

- Simple Aspect Ratio Approach: A simple method that involved using the aspect ratio of the subject's bounding box was used in [1]. This method depends on the relative position of the person, camera, and other obstacles. By mounting the camera on the ceiling was a common solution adopted by most of the researchers to overcome the occluding objects problem.

- Acoustic Monitoring: Analyzing sounds associated with falls, such as impact and distress calls, provided a passive approach. However, background noise and environmental factors impacted reliability [8,9].
- 2D Shape and Velocity Analysis and Foreground Separation and Pixel Classification : [3] detected a fall using the 2D shape and velocity of the person. [2] used foreground separation using pixel classification in order to differentiate between the subject and background.

- Pixel Classification and Color Model and Virtual Height : [5] also made use of techniques used in [2] and incorporated a color model (brightness distortion and chromaticity distortion)[4] and concepts like virtual height.

These studies demonstrate the diverse approaches and techniques employed in the field of fall detection, ranging from wearable sensors to environmental monitoring and audio-video processing.

III. Fall Detection Sytem

A. Understanding Falls

Understanding the complex process of falls in Figure 1 in adults is important for developing prevention strategies. Factors that cause falls are related to internal factors such as age regarding the ability to fall, physical ability and health, and other environmental factors. These external problems include an increased risk of falls, especially for the elderly, and environmental problems such as inadequate lighting, poor flooring, inadequate furniture or barriers.

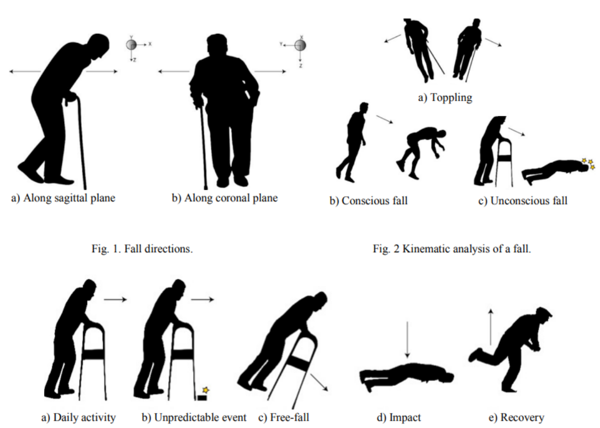

Fig. 1: Anatomy of a fall

The physical nature of falls in the elderly is associated with events that affect personal safety. This imbalance can be caused by a variety of internal factors, such as age-related loss of muscle mass, coordination, and perception. When the balance is disrupted, the body starts to work harder and this causes falls. These movements occur in two main planes: the sagittal plane (where forward or backward falls can occur) and the coronal plane (which affects both sides). During the fall, the object enters free fall, where potential energy (due to the height of the body) turns into kinetic energy as it accelerates toward the ground. This fall causes contact with the ground and may cause injury, depending on many factors such as the angle and severity of the impact, the position of the body during the fall, and the person's reaction. After a blow, there is a recovery period during which the person will try to regain balance or consciousness.

B. Extracting Landmarks

Using the media pipe landmark function, we extracted the landmarks for different body parts of a human in the feed. The coordinates for different joints are calculated in different manner -

```
# dot - LEFT_FOOT_INDEX

dot_LEFT_FOOT_INDEX_X= int(results.pose_landmarks.landmark[mp_pose.PoseLandmark.LEFT_FOOT_INDEX].x * image_width)
dot_LEFT_FOOT_INDEX_Y= int(results.pose_landmarks.landmark[mp_pose.PoseLandmark.LEFT_FOOT_INDEX].y * image_height)

# dot - RIGHT_FOOT_INDEX

dot_RIGHT_FOOT_INDEX_X= int(results.pose_landmarks.landmark[mp_pose.PoseLandmark.RIGHT_FOOT_INDEX].x * image_width)
dot_RIGHT_FOOT_INDEX_Y= int(results.pose_landmarks.landmark[mp_pose.PoseLandmark.RIGHT_FOOT_INDEX].y * image_height)
```

Fig. 2: Extracting Landmarks

C. Calculating Angles

Using the landmarks extracted from the last step we calculate the angles between different joints of human body. The method is displayed in the images below.

```
 #calculating angle

angle_shoulder_l = calculate_angle(elbow_l, shoulder_l, hip_l)
```

Fig. 3: Calculating angles using extracted landmarks

D. Extracting Points for Base of the Human Body

This part is most important as we are calculating and generating the coordinates of the base of the human body which will be one of the deciding factors of a fall. These points are denoted as points of action and the way these points of action are calculated is shown below.

```
#for feet
Point_of_action_LEFT_X = int(
    ((dot_LEFT_FOOT_INDEX_X +  dot_LEFT_HEEL_X)/2) )

Point_of_action_LEFT_Y = int(
    ((dot_LEFT_FOOT_INDEX_Y+   dot_LEFT_HEEL_Y)/2) )

Point_of_action_RIGHT_X = int(
    ((dot_RIGHT_FOOT_INDEX_X +  dot_RIGHT_HEEL_X)/2) )

Point_of_action_RIGHT_Y = int(
    ((dot_RIGHT_FOOT_INDEX_Y+   dot_RIGHT_HEEL_Y)/2) )

#co ords between feet

Point_of_action_X = int ( (Point_of_action_LEFT_X +  Point_of_action_RIGHT_X)/2 )

Point_of_action_Y = int ( (Point_of_action_LEFT_Y +  Point_of_action_RIGHT_Y)/2 )
```

Fig. 4: Points of action calculation

E. Detecting Fall

To detect the fall, we subtract the value of the x coordinate of the base (point of action) from the x coordinate of the center of mass/gravity of the body and then if it meets certain conditions it is considered a fall else the person is considered as standing. The conditions are in the image below.

```
fall = int(Point_of_action_X - dot_BODY_X )

falling = abs(fall) > 50
```

Fig. 5: Fall Condition on the basis of center of mass

IV. METHODOLOGY

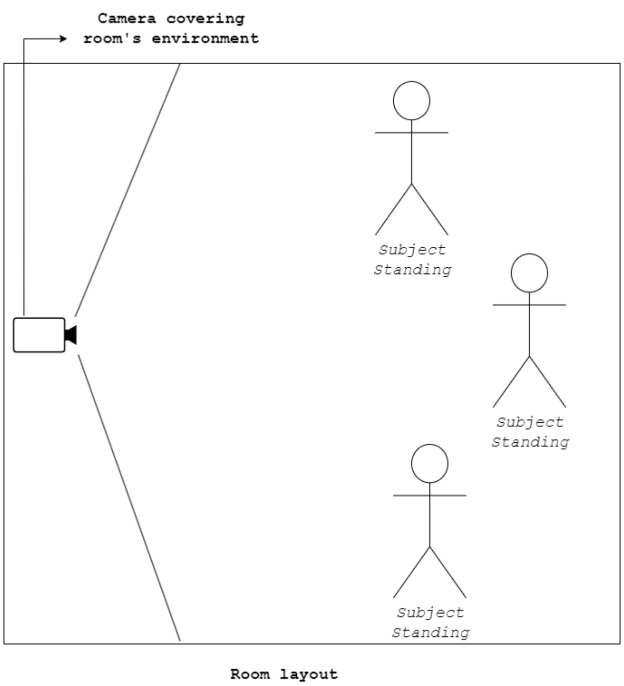

Fig. 6: Room layout before fall

A. Algorithm Used

The model uses a media pipe by Google to detect and find all the landmarks for all the major parts of the body and later using these coordinates finds the base of the human in the feed and the center of mass of that human for upper, lower and full body and then if the center of mass of the body goes out from the base of the human body, it considers it as a fall.

$$\frac{m_1 \cdot x_1 + m^2 \cdot x^2}{m_1 + m^2}$$

Fig. 7: Formula for rectangle's center of gravity

So, measuring the fall for the human body will work as follows: Human's Body center of mass * 0.75 ¿ — x-axis center of feet – x-axis center of mass of body —

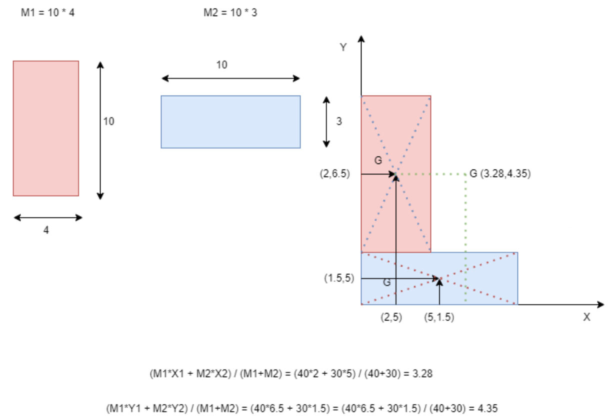

Fig. 8: Example of the center of mass of a rectangle

The fall is detected in the following method: The algorithm works in such a way that based on the coordinates detected by the media pipe framework it calculates the most probable center of mass for the human and if that condition passes then it checks for different angles of the human body parts of the person in frame and if that case passes it detects for the body of the human in frame Cnn framework that helps in differentiating between the length and breadth of the human body and if the breadth of the body is more than that of the length then, in the end, the framework considers it as a fall and sends a message to concerned people.

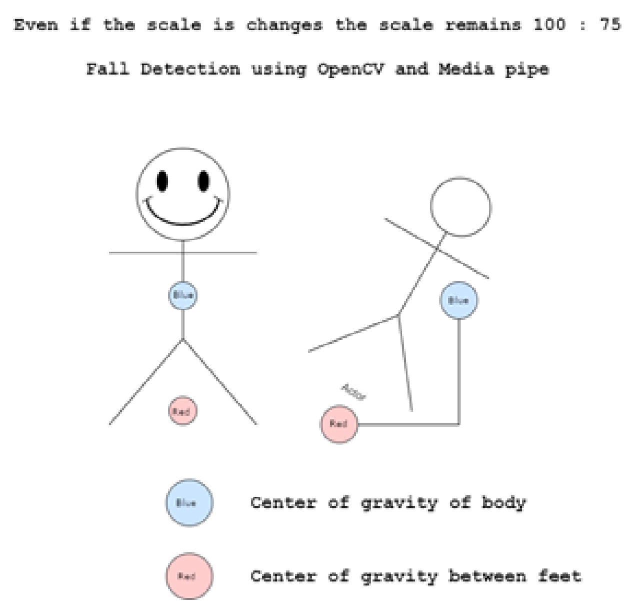

Fig. 9: Representation

B. Sound Recognition

To detect distress calls made by a person falling either within or outside the video frame, a speech recognition feature has been added. This ensures that if someone falls outside the frame and calls for help, an alert can be sent to their family. If the situation is serious, emergency arrangements can be made for help.

1) Listening to the call: To detect sound the framework keeps the microphone enabled all the time and whatever it hears is added to a string to be processed later the new string is generated every time there is an ambient pause in the audio being picked up by the microphone.

```python
def listen(image):
    with sr.Microphone() as source:
        print("Listening for 'help'...")
        recognizer.adjust_for_ambient_noise(source, duration=1)
        audio = recognizer.listen(source)

    try:
        print("Recognizing...")
        query = recognizer.recognize_google(audio).lower()
        print("You said:", query)
        return query

    except sr.UnknownValueError:
        print("Sorry, I didn't catch that.")
        return ""
```

Fig. 10: Function to enable mic and listen for help

2) Checking the condition: After a string is generated by the framework after listening to the conversation going on around it. It checks for the "help" keyword in the string and if the respective condition is met it sends a message as per the condition.

C. Message Generation:

Once a fall is detected from the video feed or a distress call using the keyword "help" is recognized, the framework sends a message through a Telegram bot. If the fall is detected from the video, an image of the fall is sent. If the distress call is detected through audio, a simple message is sent.

D. Challenges

Despite all the work done in this field in recent years, there are still some clear challenges to overcome

1) **False Alarm**: Visual fall detection is inherently prone to high levels of false positives as what appears to be a fall might not be a fall, but a deliberate movement towards the ground. In other words, most current systems are unable to discriminate between a real fall incident and an event when a person is lying or sitting down abruptly.

2) **Different Living Conditions**: All houses are not the same, different people have different lights, furniture, and floors. These things can affect how well fall detectors do their job. Making sure the system works right in all kinds of homes is a complex task.

3) **Providing Emergency Services**: For fall detection systems to work really well, they need to smoothly connect with healthcare systems. Making sure that alerts go to the right healthcare providers or emergency services on time is a challenge that involves working together with other systems.

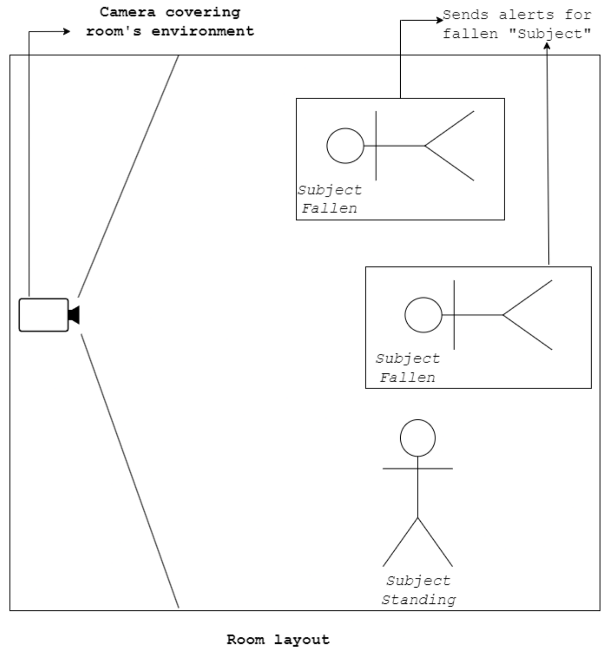

Fig. 11: Room Layout

V. Results and Discussion

A fall detection system was implemented using Media Pipe and OpenCV and tested for various positions (standing, walking, sitting). Upon detecting a fall, a message indicating the fall is displayed in the command window.

A. Fall Detection Result

In Fig. 12 the person is shown standing in light and dark. Their center of mass is within their base, their body's width is less than its length, and the inclination angle is under 45 degrees. Consequently, the output in the command window remains "no fall."

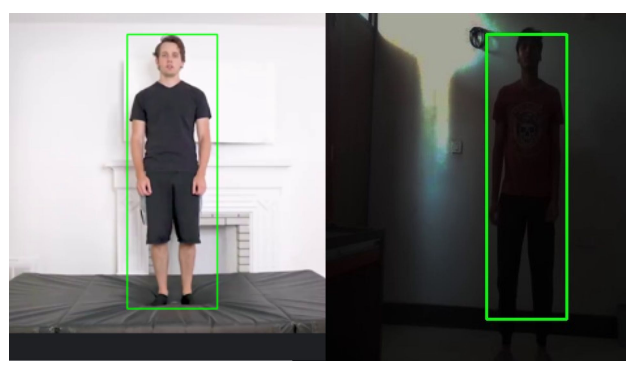

Fig. 12: Image of a person standing upright

In Fig. 13 the person is walking in daylight. All the conditions of no fall/ standing are completed and hence the output in the command window remains "no fall."

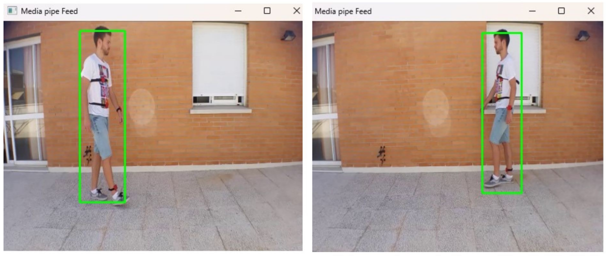

Fig. 13: Image of a person walking

In Fig. 14 the person is sitting and the center of mass of the subject is on the base of the human body. The width of the body is also less than the length. Also, the person is sitting so no false call is made for such daily activity.

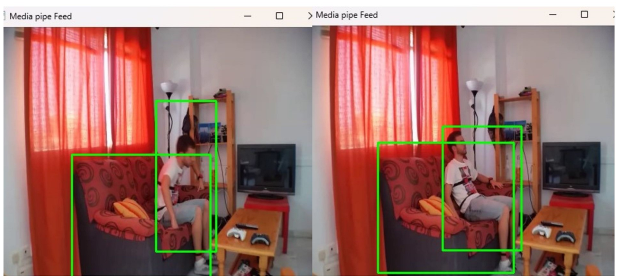

Fig. 14: Image of a person sitting

In Fig. 15 the person is falling down and the center of mass of the subject is out of the base of the human body. The width of the body is also more than the length. Also, as the person falls the message “Fall Detected!” is displayed on the screen

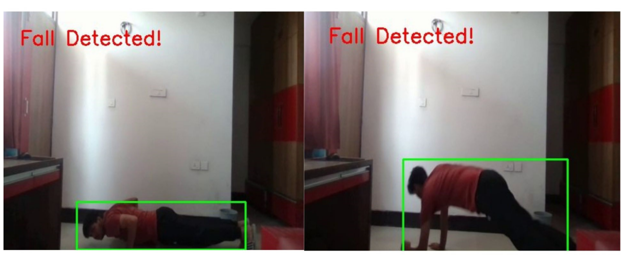

Fig. 15: Image of a person falling

B. Alert Messages Sent

As soon as the word help is detected in the speech of the subject either in or out of the frame of the camera a message is generated to send an alert to concerned parties.

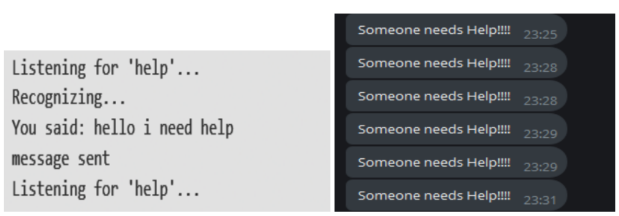

Fig. 16: Alert message generated and sent

REFERENCES

[1] B. T¨oreyin, Y. Dedeoglu, and A. C¸ etin. ”HMM based falling person detection using both audio and video.” In IEEE International Workshop on Human-Computer Interaction, Beijing, China, 2005.

[2] E. J. M.Shakeri, H.Deldari, H.Foroughi, ”A Novel Fuzzy Background Subtraction Method Based on Cellular Automata for Urban Traffic Applications”, 9th International IEEE Conference on Signal Processing, ICSP’08

[3] T. Lee and A. Mihailidis. ”An intelligent emergency response system: preliminary development and testing of automated fall detection.” Journal of telemedicine and telecare, 11(4):194–198, 2005

[4] T. Horprasert, D. Harwood, and L. Davis, “A statistical approach forreal-time robust background subtraction and shadow detection.” in Proc.IEEE International Conference on Computer Vision, ICCV99, Frame-RateWorkshop, 1999, pp. 1–19.

[5] Chen, Yie-Tarng & Lin, You-Rong & Fang, Wen-Hsien. (2012). ”Detection of Human Fall in Video Using Shadow Information.”

[6] Wolfe, Michael J.; Caguioa, Jospeh; Nguyen, Andy; and Cheun, Jacquelyn PhD (2020) ”Fall Detection: Threshold Analysis of Wrist-Worn Motion Sensor Signals,” SMU Data Science Review: Vol. 3: No. 2, Article 10. A

[7] A. Naeem and S. Khan, ”Fall Detection using Accelerometer Calibration,” *2019 15th International Conference on Emerging Technologies (ICET)*, Peshawar, Pakistan, 2019, pp. 1-6

[8] Koszalinski R, Tappen RM, Ghoraani B, Vieira ER, Marques O, Furht B. Use of Sensors for Fall Prediction in Older Persons: A Scoping Review. Comput Inform Nurs. 2023 Dec

[9] A. N. Ishak, M. H. Habaebi, S. H. Yusoff and M. R. Islam, ”Wearable Based-Sensor Fall Detection System Using Machine Learning Algorithm,” *2021 8th International Conference on Computer and Communication Engineering (ICCCE)*, Kuala Lumpur, Malaysia, 2021, pp. 53-57

[10] Shparii, Ilona, ”Real-Time Fall Detection and Response on Mobile Phones Using Machine Learning” (2017). *Master’s Theses*. 3705.