Fall Detection and Boundary Detection in Care Homes

Annette Theresa Mathew[1]; Taniya Shirley Stalin[2]; Krishna Sudheer Kumar[3]; Abhinav Santhosh[4]; Ashwin Juby[5]
[1,3,4,5]Student, [2]Assistant Professor,
Rajagiri School of Engineering and Technology

Abstract:- The elderly population represents a significant and rapidly expanding demographic, with a majority experiencing frequent daily accidents, notably falls. Falls rank as the second leading cause of accidental injury deaths globally. To address this issue, we propose a video classification system designed specifically for fall detection. Our fall detection framework comprises two key steps: firstly, the detection of human posture within video frames, followed by fall classification using Convolutional Neural Networks (CNNs). Additionally, we introduce a novel approach for boundary detection, utilizing object detection techniques beyond a predefined line of surveillance captured by a single camera. Through this integrated methodology, we aim to enhance fall detection and boundary breach detection capabilities, thereby contributing to the advancement of elderly care and safety. (Abstract).

Keywords:- *CNN; Fall Detection; Boundary Detection; Video Classification; Background Subtraction; Elders.*

I. INTRODUCTION

The elderly population continues to grow, raising concerns about the significant health impact of falls, which are a major global health concern. Recurrent falls are particularly alarming as they contribute significantly to morbidity and mortality among older adults, indicating underlying issues with physical and cognitive health. Beyond physical injury, recurrent falls can lead to psychological trauma, commonly known as "post-fall syndrome," where older adults may develop a fear of further falls and subsequent injury, resulting in a reluctance to move. While a significant portion of these falls result in minor injuries, approximately 10% lead to major injuries [1].

A fall is an unintentional occurrence where an individual ends up resting on the ground or a lower surface [2]. Falls can be categorized into three phases. The first phase involves an initial event that shifts the body's centre of mass beyond its support base, influenced by external factors or internal factors. The second phase occurs when the systems responsible for maintaining upright posture fail to detect and correct this displacement promptly, leading to the likelihood of a fall. The third phase is marked by the body's impact with nearby surfaces. This impact frequently results in the damaging transfer of force to internal tissues and organs. Subsequently, there may be medical, psychological, and healthcare consequences following the fall and associated injuries [3]. It is imperative to respond promptly to falls, as any delay can result in devastating outcomes. Though nationwide statistics are scarce, regional investigations in the US point to a serious issue of non-fracture injuries due to falls among older adults. For instance, an Ohio study demonstrated a high demand for emergency care among those 75 and older, with women experiencing these injuries at a rate of approximately 80 per 1,000 annually [4].

People with dementia frequently engage in wandering behaviour, which can pose risks and present challenges to caregivers and staff in care facilities. Traditional methods used to prevent wandering include physical restraints and medication. Dementia continues to be a significant public health concern worldwide. A 2019 report by the World Health Organization highlights the scale of the issue, with roughly 10 million new cases diagnosed each year. This alarming rate means a new person is diagnosed with dementia every few seconds[5].

In India, it is estimated that in 2020[6], 5.3 million individuals aged 60 and above are living with dementia. Incidence rates of wandering vary, with some studies reporting rates as high as 50% among individuals with severe dementia and 63% among those residing in community-based care homes[7].

Our architecture is relevant since it uses a vision-based technique to identify falls and to detect boundary breaches. For detecting falls we use a method with two stages: In the first stage we use Mediapipe to detect the human posture[8] and in the second stage we make use of Convolutional Neural Network (CNN) to classify falls. The use of vision-based fall detection system helps us in avoiding the use of wearable sensors which can be an inconvenience to the elderly population especially since they can be hesitant to use new technologies. To perform real-time boundary breach detection, object detection is performed beyond a set arbitrary line, this helps in notifying the caregivers in case a resident wanders off.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2666

II. RELATED WORKS

In this section, we will summarize the existing research on fall detection methods, specifically wearable devices and vision-based systems. Additionally, we'll examine relevant studies in the field of boundary detection systems.

Recent advancements in wearable technologies have introduced novel methods for assessing fall risk and detecting falls. Subramaniam S[9] proposed a method incorporating inertial-based sensor systems and insole-based systems for this purpose. Inertial sensors, particularly inertial measurement units (IMUs), are commonly utilized due to their ability to measure linear acceleration and angular velocity using multiple 3-axis sensors, such as accelerometers. These sensors, known for their affordability, compact size, and precision in motion-related parameter measurement, are strategically placed near the body's center to detect orientation and assess fall risk. Additionally, they can estimate posture duration and transition times between postures, with longer transitions indicating decreased balance and stability, and consequently, higher fall risk.

Conversely, insole-based sensor systems detect plantar pressure distribution and other gait variables directly from the sole of the foot, with an emphasis on gait analysis. In the past, these systems have mostly focused on measuring plantar pressure distribution, which provides important information about a person's gait and fall risk.

Wearable technologies go beyond motion tracking with sensors like those used in photoplethysmography (PPG) and electrocardiography (ECG). PPG devices detect changes in blood volume, while ECG sensors monitor the electrical activity of the heart. Additionally, spirometers[10], diagnostic tools capable of measuring lung capacity and airflow, contribute to the array of sensors available for comprehensive fall risk assessment and detection.

Despite their effectiveness in functionality, wearable devices are not preferred in aging societies due to drawbacks such as discomfort from prolonged wearing, charging issues, and the tendency to be forgotten easily. However, these limitations can be addressed by utilizing on-wearable and computer-vision-based approaches.

Computer vision-based fall detection offers advantages over wearable-based approaches due to its cost-effectiveness, non-invasiveness, and user-friendly nature. Various methodologies exist for implementing vision-based fall detection systems. One such method, proposed by Rabia Hasib[11], utilizes Mask Regional Convolutional Neural Network (Mask R-CNN) to detect and extract human silhouettes from video frames. Subsequently, the extracted human silhouette undergoes classification using Convolutional Neural Network (CNN), enabling the identification of different human postures (e.g., sit, bend, stand, lie), as well as the detection of fall events (i.e., detecting a lying posture on the ground). In a separate study conducted by Mobsite[12], a two-stage fall detection approach was introduced. In the first stage, human body silhouettes are extracted from video frames using Mask R-CNN, while the second stage involves a combination of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) to capture long-term dependencies between successive frames.

A study conducted by Prakaashini S[13] proposed an efficient and cost-effective system aimed at enhancing security measures. This system utilizes CCTV cameras for intrusion detection, leveraging the OpenCV framework. Through rigorous analysis, the system distinguishes and identifies individuals entering restricted areas. Only upon confirming the presence of a human, the system notifies the relevant authorities with a captured photo. Initially, cameras are positioned in designated no-access zones and are powered off to conserve energy. They activate solely upon detecting human presence, thus optimizing energy usage. Subsequently, the system continuously monitors these areas, promptly detecting and capturing images of humans entering restricted zones. Once an image is captured, the system automatically triggers notifications to property authorities via email, providing them with the necessary visual evidence.

III. PROPOSED METHODOLOGY

The suggested methods are explained in full in this section. The modules listed below are a part of the system architecture: The usage of object detection in a restricted region for border detection, Mediapipe for person detection on an input video, CNN for the categorization of human postures, and fall detection based on lying attitude are all covered. The general flowchart of the suggested method is shown in Figure 1.

A. Fall Detection

Figure 1 depicts the system architecture to perform fall detection on video inputs.

➢ *Video Processing and Pose Detection:*

MediaPipe is an open-source framework developed by Google that performs well in tasks like human posture recognition, hand gesture recognition and face detection. It is a pre-trained model that also supports the integration of custom models to customize according to specific use cases and domains thus enhancing the performance.

This framework is used to estimate human posture as a first step. From input video frames the image is processed using a pose estimation model. Within the processed image key-points or landmarks are drawn. These landmarks are typically joints such as shoulders, elbows, wrists or hips. Once the landmarks are drawn bounding boxes are set around the detected humans to perform further analysis.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2666

Table 1: CNN Structure Definitions of the Proposed Approach

Layers	Parameters	Values
Input	Input shape	(200, 360, 3)
Conv2D (1st layer)	Filters Kernel size Activation function	32 3x3 ReLU activation
MaxPooling2D	Pool size	2x2
Conv2D (2nd layer)	Filters Kernel size Activation function	32 3x3 ReLU activation
BatchNormalization	-	-
MaxPooling2D	Pool size	2x2
Dropout	Dropout rate	0.1
Conv2D (3rd layer)	Kernel size Filters Activation function	64 3x3 ReLU activation
BatchNormalization	-	-
Dropout	Dropout rate	0.1
Conv2D (4th layer)	Filters Kernel size Activation function	128 3x3 ReLU activation
BatchNormalization	-	-
MaxPooling2D	Pool size	2x2
Dropout	Dropout rate	0.1
Conv2D (5th layer)	Filters Kernel size Activation function	256 3x3 ReLU activation
BatchNormalization	-	-
MaxPooling2D	Pool size	2x2
Dropout	Dropout rate	0.1
Flatten	-	-
Dense (1st layer)	Units Activation function	512 ReLU activation
BatchNormalization	-	-
Dropout	Dropout rate	0.1
Dense (2nd layer)	Units Activation function	256 ReLU activation
Dense (3rd layer)	Units Activation function	128 ReLU activation
BatchNormalization	-	-
Dropout	Dropout rate	0.2
Dense (Output layer)	Units Activation function	2 Softmax activation

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2666

➢ *Convolutional Neural Network (CNN):*

Convolutional Neural Networks (CNNs) are a specialized deep learning architecture commonly applied to image analysis tasks. They use a series of convolutional layers for feature extraction, pooling layers to reduce dimensionality, and non-linear activation functions (like ReLU). Fully connected layers then perform the final classification. In fall detection applications, CNNs can be trained to recognize patterns in video data that correspond to falling postures. Our experimental neural network design is outlined in Table I.

The dataset will be divided into distinct training and testing sets. The initial layer of the CNN architecture is a convolutional layer employing 3x3 filters (32 total) to enhance feature representation. The CNN is designed sequentially, with a pooling operation following the convolutional layer. This pooling step downsamples the data, reducing its spatial dimensions. To improve efficiency, a batch normalization layer is applied after the second convolutional layer, standardizing values to a range between 0 and 1. Another max pooling layer to further downsample the feature maps. The next three sequences of convolutional layers consists of 64, 128 and 256 filters respectively. Each of these convolutional layers is followed by a pooling layer. Dropout layers are used to prevent overfitting by randomly dropping random values during training. Later in the flatten layer 2D output is converted into a 1D vector to feed it into the fully connected layers. This neural network model uses a three-layer, fully connected architecture. The hidden layers have 512, 256, and 128 units respectively, and use ReLU activation. The output layer has two units with softmax activation, designed for binary classification tasks.

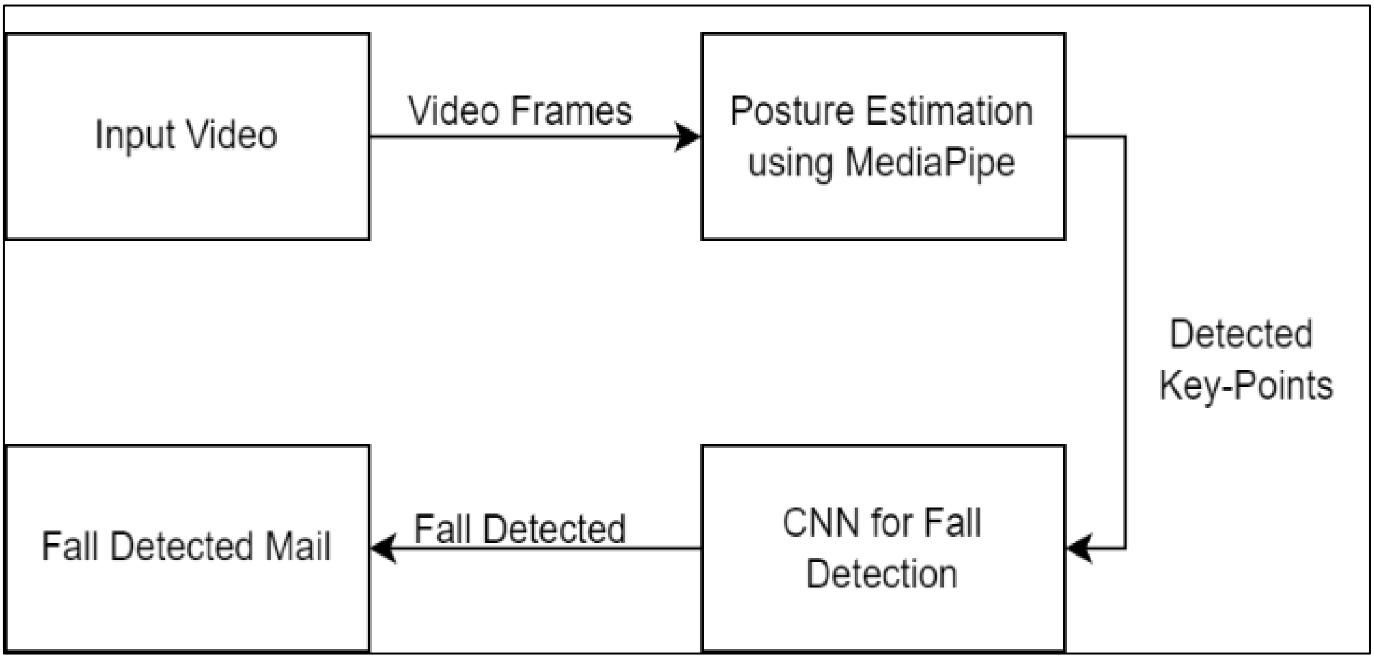

Fig 1: System Architecture of Fall Detection System

B. Boundary Breach Detection

Figure 2 depicts the system architecture to perform boundary breach detection on live video inputs.

➢ *Background Subtraction and Contour Detection:*

Background subtraction is a technique that is applied in order to separate moving bodies on stationary background. It is particularly useful for tasks where the focus is on detecting changes in the scene over time, such as surveillance. It helps in eliminating static or stationary elements. For tasks such as boundary detection, it helps in identifying moving humans while the background remains static. The result of background subtraction is a foreground mask that consists of a moving object against a black background. Contour detection is then performed on the extracted foreground. This helps in delineating the boundaries of the extracted foreground. By detecting contours in the foreground mask, we can precisely locate the boundaries of the moving objects, enabling further analysis. Contour detection provides information which is essential for tasks such as object recognition and tracking.

➢ *Trespass Detection:*

Trespass detection is the analysis of detected contours to determine if an object has crossed a predefined boundary. The input for this module would be the detected contour. The area of the contour in calculated. If the area of the contour exceeds a predefined threshold, further processing is done to determine if the contour represents a trespassing object. Once the area is found, the centroid of the detected contour is detected and its position is compared with the predefined boundary. If the centroid lies inside the boundary, it indicates trespassing. This detection triggers alerts to relevant authorities, enabling timely response to potential boundary breaches.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2666

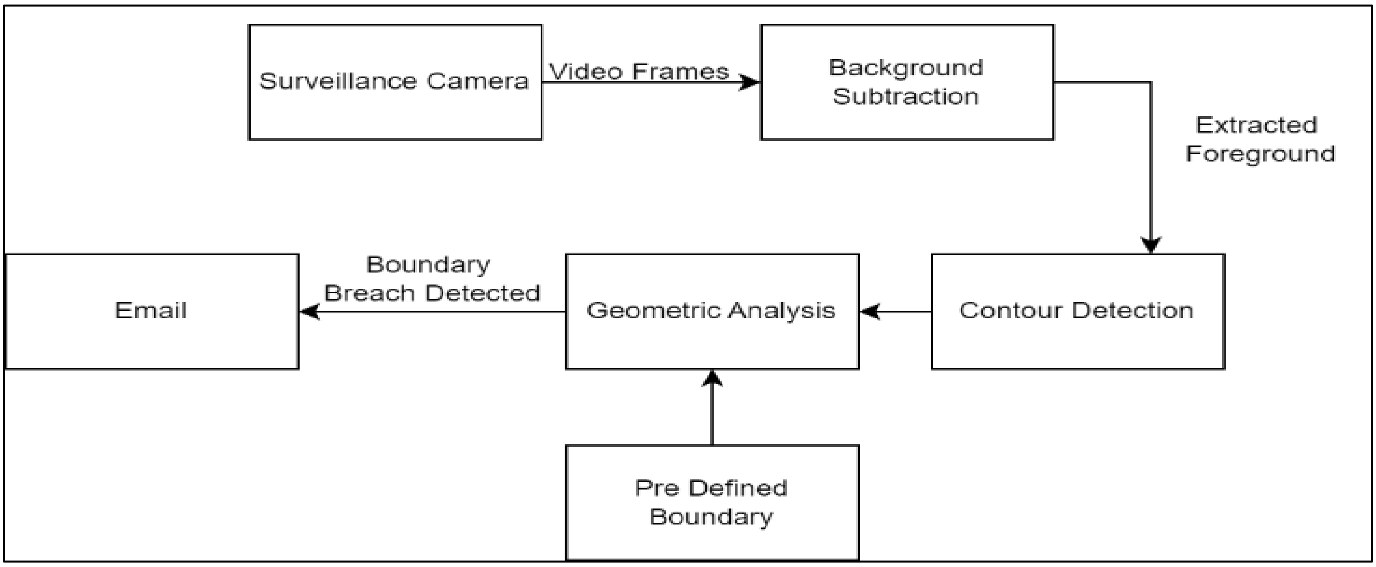

Fig 2: System Architecture of Boundary Breach Detection System

IV. RESULTS AND DISCUSSION

Testing occurred on a system equipped with an Intel Core i7 processor and 8 GB of RAM. We leveraged the Google Colab environment for development, employing Python and the Tensorflow framework for the core implementation. To assess performance, we used a proprietary dataset with an 80:20 split for training and testing purposes.

The classification of human postures and the detection of boundary breaches is performed in the context of old age homes or care homes. For fall detection, the input videos are converted into frames of size (360,200) and it is passed through a sequence of convolutional layers, pooling layers and later through fully connected layers in order to perform binary classification and detect falls accurately. In the case of detection of boundary breaches, background subtraction is performed to create a separatio between a moving body and a stationary background. The proposed achieves an accuracy of .791 and a loss of .650.

A. Performance Metrics

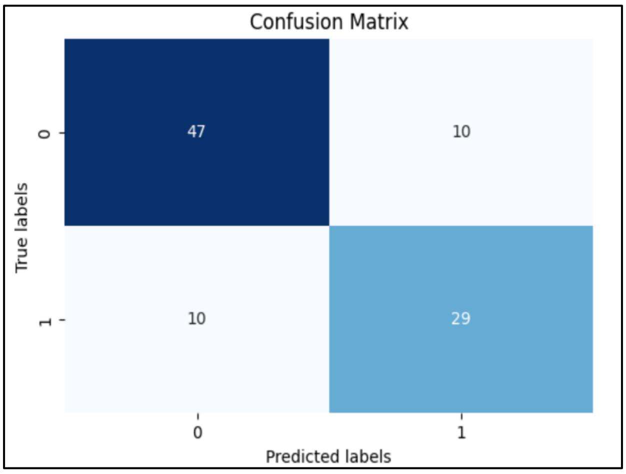

Fig 3: Confusion Matrix

The confusion matrix, shown in Figure 3, is used to assess classification accuracy and assess the overall effectiveness of the model. TP, TN, FN, and FP are the parameters that are utilized to calculate the confusion matrix. TP known as true positive denotes the instances that have been correctly identified as fall by the model whereas FP or false positive denotes the instances that have been wrongly identified as fall. TN known as true negative denotes the instances that have been correctly identified as not fall by the model whereas FN or false negative denotes the instances that have been wrongly identified as not fall. The models capabilities can be assessed using accuracy, specificity, sensitivity or recall and F1 score using the equations from (1) to (4).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Sensitivity\ or\ Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1\ score = 2.\left(\frac{Precision * Recall}{Precision + Recall}\right) \tag{4}$$

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2666

B. Accuracy and Loss Plots

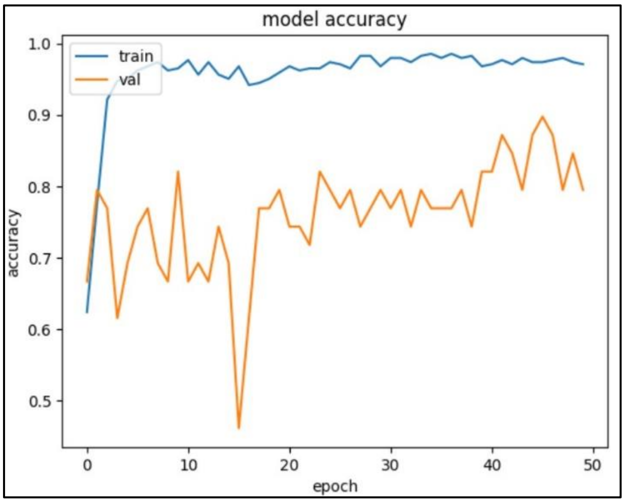

Fig 4: Accuracy Graph

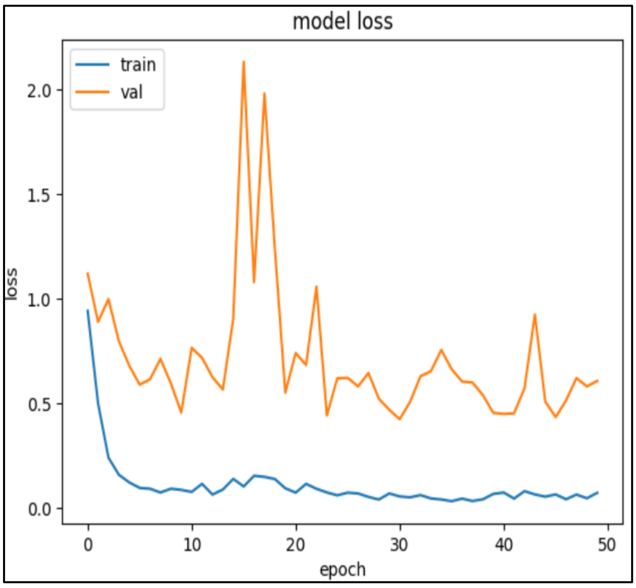

Fig 5: Loss Graph

Figure 4 depicts the progression of training and validation accuracy across epochs. The training accuracy curve reveals how well the model learns to predict the training data as training progresses. The validation accuracy curve reflects the model's performance on an unseen dataset. Analysing the relationship between these curves allows us to identify potential issues like overfitting or underfitting.

Figure 5 presents the evolution of training and validation loss over epochs. The loss plot demonstrates how effectively the model reduces errors during the training process. As with accuracy curves, comparing training and validation loss curves aids in diagnosing overfitting or underfitting tendencies.

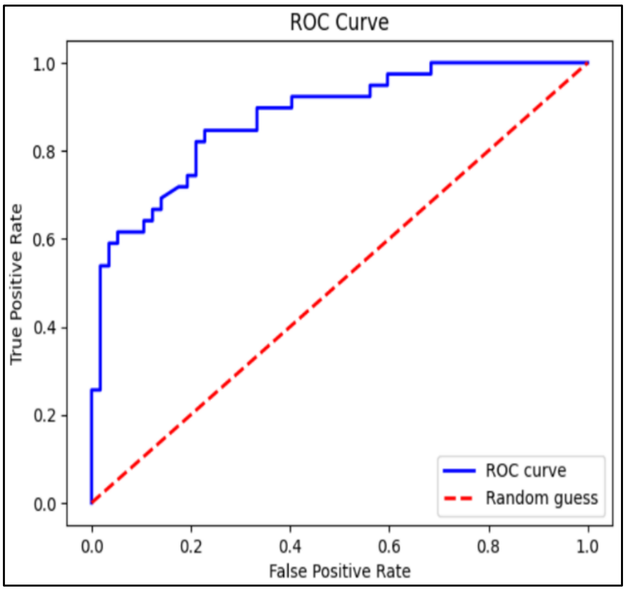

Fig 6: ROC Curve

To thoroughly assess model performance, we utilized the Receiver Operating Characteristic (ROC) curve and its related metric, the Area Under the Curve (AUC). Refer to Figure 6 for a visual representation. The ROC curve reveals the balance between a model's ability to correctly identify violent incidents (True Positive Rate - TPR) and its rate of mislabeling non-violent incidents (False Positive Rate - FPR). The AUC metric condenses this information into a single value, where a higher AUC signifies superior classification performance. Examining the ROC curves and AUC values allows us to evaluate the diagnostic potential of each model, aiding in determining their suitability for practical violence detection scenarios.

V. CONCLUSION

For fall detection, the study employs a vision-based approach leveraging techniques such as human posture detection using Mediapipe and Convolutional Neural Network (CNN) classification. Mediapipe offers an effective pose estimation capabilities, enabling the extraction of precise human body landmarks from video frames, while CNNs provide efficient classification of fall patterns based on extracted features. Additionally, boundary breach detection utilizes background subtraction algorithms for foreground extraction, contour detection to delineate moving objects, and centroid analysis to determine object positions relative to predefined boundaries. These techniques, coupled with thresholding and comparison operations, enable accurate detection of trespassing incidents. Furthermore, the integration of alerting systems enhances the responsiveness of the overall detection framework, ensuring timely intervention in critical situations.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2666

REFERENCES

[1]. Vaishya R, Vaish A. Falls in Older Adults are Serious. Indian J Orthop. 2020 Jan 24;54(1):69-74. doi: 10.1007/s43465-019-00037-x. PMID: 32257019; PMCID: PMC7093636.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2]. Kellogg International Work Group on the Prevention of Falls by the Elderly. The prevention of falls in later life. Danish Medical Bulletin 1987; 34(4):1-24.

[3]. Institute of Medicine (US) Division of Health Promotion and Disease Prevention; Berg RL, Cassells JS, editors. The Second Fifty Years: Promoting Health and Preventing Disability. Washington (DC): National Academies Press (US); 1992. 15, Falls in Older Persons: Risk Factors and Prevention.

[4]. Fife, D., Barancik, J. I., and Chatterjee, M. S. Northeastern Ohio trauma study. II. Injury rates by age, sex and cause. American Journal of Public Health 1984; 74(5):473-478.

[5]. Price JD, Hermans DG, Grimley Evans J. Subjective barriers to prevent wandering of cognitively impaired people. Cochrane Database Syst Rev. 2000;2001(4):CD001932. doi: 10.1002/14651858.CD001932. PMID: 11034735; PMCID: PMC8406984. K. Elissa, “Title of paper if known,” unpublished.

[6]. Kumar CTS, Shaji KS, and Varghese MNM (Eds.). *Dementia in India*, 2020, Cochin Chapter. Published online 2020. 1–96. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7]. Hope T, Tilling KM, Gedling K, et al. The structure of wandering in dementia. *Int J Geriatr Psychiatry*, 1994; 9 149–155.

[8]. Garg, Shubham & Saxena, Aman & Gupta, Richa. (2022). Yoga pose classification: a CNN and MediaPipe inspired deep learning approach for real-world application. Journal of Ambient Intelligence and Humanized Computing. 14. 10.1007/s12652-022-03910-0.

[9]. Subramaniam S, Faisal AI and Deen MJ (2022) Wearable Sensor Systems for Fall Risk Assessment: A Review. Front. Digit. Health 4:921506. doi: 10.3389/fdgth.2022.921506

[10]. Karar ME, Shehata HI, Reyad O. A Survey of IoT-based fall detection for aiding elderly care: sensors, methods, challenges and future trends. *Appl Sci.* (2022) 12:3276. doi: 10.3390/app12073276.

[11]. R. Hasib, K. N. Khan, M. Yu and M. S. Khan, "Vision-based Human Posture Classification and Fall Detection using Convolutional Neural Network," 2021 International Conference on Artificial Intelligence (ICAI), Islamabad, Pakistan, 2021, pp. 74-79, doi: 10.1109/ICAI52203.2021.9445263

[12]. S. Mobsite, N. Alaoui and M. Boulmalf, "A framework for elders fall detection using deep learning," 2020 6th IEEE Congress on Information Science and Technology (CiSt), Agadir - Essaouira, Morocco, 2020, pp. 69-74, doi: 10.1109/CiSt49399.2021.9357184

[13]. e-ISSN: 2582-5208 International Research Journal of Modernization in Engineering Technology and Science Volume:02/Issue:09/September-2020 Impact Factor- 5.354 www.irjmets.com www.irjmets.com @International Research Journal of Modernization in Engineering, Technology and Science.