

An Image Based Fall-Detection System for the Elderly using YOLO algorithm

Mounika Pokala

School of Computer Science and
Engineering

Vellore Institute of Technology
Chennai, India

mounika.pokala2019@vitstudent.ac.in

Abishek S

School of Computer Science and
Engineering

Vellore Institute of Technology
Chennai, India

abishek.s2019a@vitstudent.ac.in

Sriram Namana

School of Computer Science and
Engineering

Vellore Institute of Technology
Chennai, India

sriram.namana2019@vitstudent.ac.in

Abstract—Falling is one of the most frightening experiences that older people might experience, and it can also prove to be fatal at times. As the 'elderly population' grows, so does the demand for a fall detection system that can inform the system administrators in the event of a fall. The detection of a fall and early notifications will guarantee that adequate care is given as soon as possible, and that fatal injuries are averted as a result of the quick response.

To implement the 'Fall Detection' module, we will be using the YOLO algorithm to detect the incident. In this module, we take the input video from a source and divide the video into several frames. If the fall is detected continuously in 20 frames of the video, then an alert is sent to the system admin. The alert is in the form of an email which includes the detected picture.

Keywords- Elderly, Fall detection, YOLO

I. INTRODUCTION

Due to the ever-increasing global trend of single child policy and advanced applications of tech in the medical field, many countries are heading towards an aged population. Hence, one of the most important issues to tackle in the medical field is that of providing world-class elderly care facilities. According to a report from the World Health Organization, falls are the second leading cause of unintentional injury deaths worldwide. 37.3 million falls that are severe enough to require medical attention occur each year. An estimated 684,000 fatalities from falls occur each year globally of which over 80% are from low- and middle-income countries. The majority of these falls occur in citizens who are over 60 years of age.

A lot of works have already been proposed to try and tackle this problem. However, there are limitations that exist in each method. Hence, we propose a unique image-based FD system which can immediately notify the required people via an email alert that a fall has occurred. So far, we can broadly classify the works that have been already done in this regard into 2 categories: sensor-based and image-based models.

For sensor-based implementations, we mostly find solutions in the form of wearable devices. A highly accurate, low-cost machine-learning based FD algorithm [3] was proposed which makes the use of wearable sensors. A system for

elderly person monitoring [2] that is based on smart sensors worn on the body and operating through consumer home networks was also proposed. The information required for detecting the fall is collected from a combination of accelerometer, cardiometer and smart sensors. However, the main drawback with a wearable sensor-based implementation is that most elderly citizens tend to forget to wear such devices.

Image-based implementations have attracted more interest and are generally more efficient in terms of accuracy of detecting falls. For example, a video surveillance system [1] was proposed to monitor falls in elderly citizens. The same is done by performing real-time feature extraction using MLP neural network on the posture of the person. A new system that is based on the depth images captured by Microsoft Kinect [18] was also put forward. On these images skeleton tracking and bounding box analysis are done to detect the fall.

Image based methods provide several advantages over the existing wearable device based methods like, ease of use for the elders and the need to not charge the device frequently. It is difficult for the elderly people to maintain such a fragile device. As there are no physical sensors used in the image-based method, the usage of such devices becomes easier for the elderly.

There are several types of challenges that the design of FD systems faces which are explained below.

1) Performance under real-life conditions:

Fall detection systems need to be as accurate and reliable as possible. The robust fall detection model being trained to use for elderly people, it is more important for the system developed to exhibit high sensitivity and specificity and accuracy. This is sometimes reached in experimental environments, but when applied to a real situation, the detection rate decreases.

Also, only few studies incorporate data from older people, since their participation is limited to perform a set of simulated activities of daily living for few hours.

2) Acceptance:

Little is published about the practicality and acceptability of the technology. Elders' acceptance poses a major problem since they may not be familiar with electronic devices. To overcome this challenge, the way the system operates is essential. The detector should activate and operate automatically, without user intervention.

Numerous methods have been developed to detect falls in the elderly that were subjected to various challenges. With the wearable device innovation, the authors provided a review of the approaches that have been proposed for fall assessment, prevention, and detection that use wearable devices and implemented methods that adopted triaxial accelerometers to collect fall signals. Under the environmental sensors innovations. The authors used a smart floor that was embedded with pressure-sensitive fibre sensors to detect fall events by feature-specific pressure images containing motion features for human activity analysis.

In this paper, we propose an image-based FD algorithm to detect falls in the elderly. The same is done with the use of YOLO algorithm.

II. RELATED WORKS

Yacchirema et al. [4] proposed an IoTE-Fall system, an intelligent system for detecting falls of elderly people in indoor environments which uses IoT and ensemble machine learning algorithms. IoTE-Fall system employs a 3D-axis accelerometer embedded into a 6LoWPAN wearable device capable of capturing in real time the data of the movements of elderly volunteers. The acceleration readings are processed and analyzed at the edge of the network using an ensemble-based predictor model that is identified as the most suitable predictor for FD.

Saleh et al. [5] provided a review which proposes an efficient machine learning-based algorithm was proposed for wearable fall detectors. It is based on describing the physical activity of the elderly using local binary features which are built using acceleration thresholding. The system enjoys an attractive property that the computational cost of decision-making is independent from the complexity of the trained machine.

Khraief et al. [6] proposed weighted multi-stream deep convolutional neural networks that exploit the rich multimodal data provided by RGB-D cameras. The method automatically detects fall events and sends a help request to the caregivers. The contribution is three-fold. First, a new architecture was built, composed of four separate CNN streams, one for each modality. The second contribution is the combination of the four streams to generate the final decision for FD. The third contribution is the application of transfer learning and data augmentation to increase the amount of training data, avoid overfitting and improve the accuracy.

Zhang et al. [7] used a approach to enforce the temporal stability in low light video enhancement with only static images. The key idea is to learn and infer motion field (optical flow) from a single image and synthesize short range video sequences. This method can infer motion prior for single image low light video enhancement and enforce temporal consistency. The overall process being performed is low light video enhancement with image-based model and alleviating flickering by temporally stabilizing it. With the help of generated optical flow, the model is guided to learn temporal stability by enforcing consistency on warped outputs.

Palaua et al. [8] offers a unique IoT-based system for detecting older people's falls in interior environments, which utilises low-power wireless sensor networks, smart devices, big data, and cloud computing. First, a 3D-axis accelerometer placed in a 6LoWPAN device wearable collects data from old people's movements in an interior environment. Second, a decision trees-based Big Data model that is constructed and trained in the cloud detects an elder's fall. It is initially trained using historical data from an open dataset collecting records of elderly people's falls and activities of daily living. The model then learns about the fall events that the system has detected. One of the model's primary breakthroughs is that it runs on a resource-constrained device, a Smart IoT gateway with fog computing characteristics that enable fall-detection-related processing to be done locally, reducing "long lie" time. When a fall is detected, the Smart IoT gateway can send warnings to healthcare experts through a lightweight and secure IoT protocol, including information on the type of fall and the location of the old person's home. The Smart IoT Gateway also provides interoperability and data transformation, allowing the system to interface with other AAL systems or IoT Platforms in a holistic manner. Finally, the identified fall incidents are recorded in the cloud to provide healthcare experts with more precise data. Additionally, every time a fall is detected, this data is used to create a new model, which is then instanced at the Smart IoT Gateway. In terms of accuracy, precision, and gain, the FD-performance system's has been proven.

Badgujar et al. [9] proposed variety of activities of multiple participants to calculate features. Machine learning algorithms SVM and decision tree are used to detect the falls on the basis of calculated features. The system acquires accuracy up to 96% by using decision tree algorithm. Wearable sensor based FD systems are more suitable for elderly people because it can detect the fall any time and any place unlike vision based and ambient based FD which are restricted to the house or particular indoor environment. Also wearable sensor is lesser in cost than that of camera or PIR sensors. Too many sensors also lead to miss prediction of the accurate activity. Fall can be detected using two techniques after collection of data from sensors and feature calculation. One is threshold based, if reading of the sensor is above particular threshold, it can be categorized as fall. In this technique, threshold for each calculated feature is different and many false alarms may be generated. If machine learning classification is used, the calculated features can be tested on a pre-trained model with high accuracy and guarantee lesser false alarms. In the current work, accelerometer data from a wearable sensor is used, which is already measured for different activities in SisFall dataset. Most relevant features are calculated. Machine learning models of SVM and decision tree are trained and tested. In literature, these features are used for threshold based FD and the proposed work implements machine learning approach on same features and provides improved accuracy.

Santiago et al. [10] presents a FD system that monitors in real-time an older adult. The system defines two major components: a wearable device and a cell phone. The wearable has the capability of communicating with a cell phone can be located in a 100ft radius. Once, the wearable device detects a fall, it sends an alert to the cell phone; then

the cell phone alerts to the emergency contacts defined by the user. In addition, our system has a panic button that can be used in order to alert the emergency contacts in the event that the user feels that a fall may happen.

Sowmyayani et al. [11] suggested solution to introduce keyframe-based FD. The University of Rzeszow (UR) Fall Detection dataset, Fall Detection Dataset, and MultiCam dataset were used in the experiments. The suggested method yields higher accuracy rates of 99 percent, 98.15 percent, and 99 percent, respectively, for the UR Fall detection dataset, Fall Detection Dataset, and MultiCam dataset. When the proposed method's performance is compared to that of other approaches, it is found to have a greater accuracy rate. Video sequences are segmented into Groups of Pictures (GOP) based on scene changes in the suggested method. In each GOP, the keyframes are introduced. To discover big motion between GOPs and the occurrence of fall within a GOP, optical flow is used. The suggested method's performance is compared to that of recent approaches, and it proved to be superior by achieving a 99 percent accuracy in 5 seconds for UR FDD.

Sumiyaa et al. [12] suggested a mobile robot that can detect and record falls in the elderly for the aim of life-care. The robot follows the elderly from a set distance in the house, detects falls in real time, and sends e-mails to observers to alert them to the fall if it occurs. The family, helpers, and rescuers may all contact each other right away thanks to the notification. A household mobile robot (Yujin Robot's Kobuki), a sensor (Microsoft's Kinect), and a computer (PC) to detect a target and control the robot make up the mobile robot. In FD, PC software gathers bone data from the detected human using Kinect and calculates the fall using the discrepancies between the target's head and knees. The sensor is put on the robot to follow the target harmony for robot simplicity and accurate FD. As a result, the sensor can move around with the robot to reduce the amount of blind region. In comparison to a traditional monitoring technique using position-fixed sensors, the results of our trials demonstrate an improvement of up to 80% in FD rate. In their preliminary trial, they found it impossible to detect falls in 360 degrees with a fixed sensor; however, by introducing a feature that actively moves the sensor (Active Sensing), they were able to enhance the detection rate by 80%. They discovered, however, that the approach for re-detection of a user who had already fallen was insufficient, and that there are directions in which falls cannot be detected due to individual variances in robots, and that addressing these difficulties is critical.

Delahoz et al. [13] surveys the state of the art in FD and FP systems, including qualitative comparisons among various studies. It aims to serve as a point of reference for future research on the mentioned systems. A general description of FD and FP systems is provided, including the different types of sensors used in both approaches. A 3-level taxonomy associated with the risk factors of a fall is proposed. Finally, cutting edge FD and FP systems are thoroughly reviewed and qualitatively compared, in terms of design issues and other parameters.

Zhong et al. [14] proposed a wearable device-based FD system. The tech tracks the movements of the human body, detects a deviation from regular everyday activities using a powerful quaternion algorithm, and instantly alerts caretakers to the patient's location. The device's hardware and software are primarily based on a single triaxial accelerometer and GPS/GSM module. The device has an efficient FD algorithm that consumes minimal resources and power, indicating that it is well-suited for outdoor use. Because the algorithm does not require the axes of the accelerometer to be fixed tightly, there are no special requirements for the device's mounting orientation. The system has a low-power hardware design and a highly efficient algorithm that could increase the wearable device's service life. Wearable and outdoor applications are possible with both the hardware and software designs.

Kwolek et al. [15] showed how to develop and deploy a low-cost FD system with a low false alarm rate. On the basis of accelerometric data and depth maps, the fall is detected. A tri-axial accelerometer is utilised to detect the possibility of a fall as well as whether or not the individual is moving. The programme extracts the person, calculates the characteristics, and then executes the SVM-based classifier to authenticate the fall alarm if the measured acceleration is greater than an anticipated threshold value. It is a 365/7/24 embedded tech that allows for unobtrusive FD while maintaining the user's privacy. Both the Kinect sensor and a wearable motion-sensing gadget are used in this system. Our method can consistently distinguish between falls and activities of daily life when both devices are employed. The amount of false alerts is reduced when the system is configured in this way. Visual assessment of the fall alert generated only on the basis of motion data results in a lower proportion of false alarms. The authenticity of the alert is based on depth data and the analysis of features collected from depth maps.

Wadmare et al. [16] developed a UR dataset of kinetic sensor's output videos is used which classify into two categories i.e. fall and activities of daily living which are considered for experimentation. The MOG2 algorithm is used for background subtraction to focus on the human object and ignore the rest of the surrounding environment. The Shi Tomasi algorithm is used for finding interest points that are tracked using the optical flow method: Lucas-Kanade algorithm. Using interest points we compute its maximum displacement along with direction and speed of motion. For fall activity the last positions of interest points were observed. So the system detects a person in the environment, keeps tracking the person, and calculates the optical flow in case of a fall.

Nooruddin et al. [17] research offers a centralised, unobtrusive Internet of Things (IoT) based device-type invariant FD and rescue system for real-time surveillance of a large population. In the suggested system, any sort of device, like as smartphones, Raspberry Pis, Arduinos, NodeMcus, and Custom Embedded Systems, can be utilised to monitor a big population. The devices are stowed in the left or right pant pocket of the user. The accelerometer data from the devices is continuously transferred to a multithreaded server, which hosts a machine learning model that analyses the data to determine whether or not a fall has occurred. The classification results are sent back to the devices by the

server. If a fall is detected, the server sends an SMS to the mediator informing them of the user's location. The corresponding device, as a failsafe, sounds the buzzer to notify nearby folks and notifies emergency medical services and mediators through SMS for quick medical aid, thus saving the user's life. Finally, because no external connections are required, the suggested system can be deployed on a variety of devices and used to reliably monitor a large population with a low false alarm rate without interfering with the users' daily lives.

Alsadoon et al. [18] proposed a new system that is based on the depth images captured by Microsoft Kinect, skeleton tracking and bounding box analysis. The key novelty of this system is that it wraps the moving object into the bounding box and determines the change of size of the moving object by analysing the motion over the time to distinguish the human moving object and non-human moving object. The system stores the joint measurements of the known people in a database and compares the joint measurements of the detected person with the values in the database to identify the person.

Agrawal et al. [19] described a visual surveillance system for detecting human falls in this paper. To discover the foreground items, the initial step is to do background subtraction using Improved GMM. In the second stage, the human or nonhuman item is classified using contour based human template matching. It aids in the detection of falls by displaying a sudden shift in the produced score after matching. In the third stage, the height-width ratio is calculated to determine if the human shape has changed or not. The distance between the top and mid-center of the rectangle covering the human is calculated in the fourth phase, and if it is less than a specified threshold, the human fall is confirmed. Finally, if a human remains inactive for 100 consecutive frames, an alarm is triggered to warn those at home to administer treatment as soon as possible.

Lasheng et al. [20] examined the peculiarities of human fall behaviour in order to develop a FD system. By examining the internal correlation between sensor data, a long-term and short-term memory recurrent neural network is employed to increase the effect of falling behaviour detection. The first step is to build a serialisation representation mechanism for sensor data, training data, and detection input data. The BiLSTM network is used to minimise the dimension of the data required by the FD model. It has a good capacity to sequence modelling. The BiLSTM FD training method and the BiLSTM-based FD algorithm then translate the FD issue into a classification problem for the input sequence. Finally, the TensorFlow platform was used to develop the BiLSTM-based FD system. A bionic experiment data set that simulates a fall was used to detect and analyse the system.

Khawandi et al. [21] provided solution for the critical problems that are associated with those solutions like button is often unreachable after the fall, wearable devices produce many false alarms and old people tend to forget wearing them frequently. To solve these problems, Here they propose an automated monitoring that will detect the face of the person, extract features such as speed and determines if a human fall

has occurred. An alarm is triggered immediately upon detection of a fall.

A. Abbreviations and Acronyms

FD – Fall Detection

Tech – Technology

III. PROPOSED METHODOLOGY

The methodology of the object detector and distance estimator works as a subdivided module, namely the detector part which primarily uses the YOLO V4. Yolo stands for You only look once and the basis of it is that it's an object classification algorithm. The model, due to its high predictive speed and accuracy and impeccable learning capabilities happens to be perfect for object detection.

This helps the model work in real-time to detect objects in the frame of videos or images. The algorithm processes the frame of images and creates residual blocks of different sizes and hence uses the weights that it is trained on to classify those blocks into detected object,



Fig 1. Frames of the proposed object detection model

Therefore, once the model detects our frames, it uses the model to predict the object in the frame, since it is highly accurate, it classifies all the possible objects it can see in the camera thus we set a threshold for it to only detect persons in the image. The next step is to essentially calculate the distance between the detected object and the camera. Since we have an unavailability of a depth camera – used for professional depth estimation for self-driving cars, we use an approximation function to calculate the distance using only a single web camera. The way to do this is to save reference images of the object's calculated distance from the camera along with the width of the object, we then save it into variables for later reference.

```
KNOWN_DISTANCE = 30
PERSON_WIDTH = 15
```

For every detected object we call in a function which estimates the distance based on the formula below:

```

def focal_length_finder(measured_distance, real_width, width_in_rf):
    focal_length = (width_in_rf * measured_distance) / real_width

    return focal_length

def distance_finder(focal_length, real_object_width, width_in_frm):
    distance = (real_object_width * focal_length) / width_in_frm

    return distance

```

The distance, although not highly accurate, suffices as a prototype and creates a good approximation of the users distance from the object. Lastly, we move on to calling a synchronous function that provides an audio feedback of the distance and the name of the object as long as it is lesser than 20 inches from the user and the feedback is provided every even interval of inches.

The Fall detection module in our project has 2 main functionalities, to detect the fall of an individual and to send the alerts in the form of an e-mail, in an event of a fall to the concerned individuals. To collect the data from a given environment or surrounding, we use a video recording camera with a suitable frame rate and picture quality. On the data collected from the camera we perform the object detection, to detect the fall of an individual. In order to perform the object detection on a real – time basis, we have chosen to use the YOLOv4 algorithm.

The main goal of using the YOLOv4 algorithm is to detect the presence of an individual on any part of the frame. The model is trained on the **COCO** dataset, where we use only the **person** class to train the model. This is because we want our model to be able to detect and segment only the humans present in the frame. Once the video is captured by our device it is then divided into multiple frames, where each frame goes through a set of pre-processing steps which is done by the OpenCV library. We use this library to draw bounding boxes of our target in the image. Once a human is detected in the frame, the image of the individual on the frame is then segmented by drawing bounding boxes around the region of the individual in the image.

Once the bounding boxes are drawn, we will be able to check the corresponding height and width of the individual in the image. Once the width of the bounding box exceeds the height of the bounding box, we classify this scenario as a fall. But the alert/email is not sent immediately after a fall is detected. We wait for the classification of a fall to take place continuously for at least 20 frames. This number of frames that is taken as a threshold/trigger can vary based on the video recording device chosen for a given scenario.

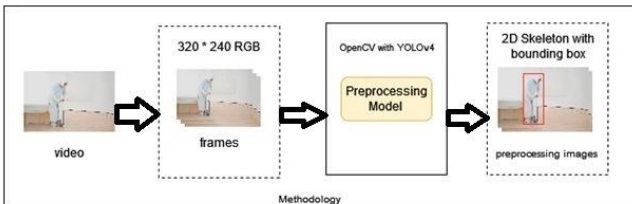


Fig 2. Architecture of the Fall Detection model

Once the number of classifications crosses the threshold, an email/alert is sent by using the smtp protocol. In order to implement the alerting mechanism of our work, we use the

smtp python library. This python library can be used for sending mails and handling the routing between mail servers. In our work, in order to implement an alerting mechanism to get attention as soon as possible, we use this library to define and create a client session in order to send a mail to the concerned individuals. The contents of the mail that is sent consists of the image of the detected fall of the individual.

IV. PERFORMANCE ANALYSIS

Real time object detection was implemented using the YOLO V4 algorithm. YOLO V4 was opted for in this module as the aforementioned is an enhanced version of its predecessors, the YOLO and YOLO V2 algorithms in terms of accuracy, speed and architecture.

YOLO V2 uses Darknet-19 as its primary feature extractor. Darknet-19 is a convolutional neural network, which, as the name suggests, is 19 layers deep. YOLO V4, as an improvement uses Darknet-53 for the same, a 53 layer deep CNN making it a much more efficient algorithm, even beating the well-known ResNet-101 and ResNet152.

Table 1. Comparison of performance measures of the YOLO algorithms

Backbone	Top-1	Top-5	Ops	BFLOP/s	FPS
Darknet-19	74.1	91.8	7.29	1246	171
ResNet-101	77.1	93.7	19.7	1039	53
ResNet-152	77.6	93.8	29.4	1090	37
Darknet-53	77.2	93.8	18.7	1457	78

From Table 1, we can see that Darknet-52 is 1.5 times quicker than ResNet101 using the chart provided in Redmon and Farhadi's YOLOv4 publication. Because it is still as precise as ResNet-152 while being two times faster, the shown accuracy does not imply any trade-off between accuracy and speed between Darknet backbones.

The Fall Detection system started with input from the camera feed where the video frames were captured in real time. If a fall has been detected in twenty continuous frames, then an alert is sent to the user. The YOLO algorithm is used to detect the person/object and OpenCV is used to draw the bounding boxes around the detected object.

The system accuracy was evaluated based on different test cases and scenarios. The different test cases used to measure accuracy were:

a. Fall events which occurred while sitting under daylight test case

Table 1 depicts the result for the above-mentioned test case. About 30 fall scenes were evaluated under broad daylight. For the 30 fall scenes, mail was sent for 29 scenes by the fall

detection system. An accuracy of 96.7% was obtained for this test case.

Table 2. Fall events evaluated under broad daylight while sitting

Test Cases	Actual Fall	Detected Fall	Accuracy
Total Cases	30	29	96.7

b. Fall events which occurred while standing under daylight test case

Table 2 depicts the result for the above-mentioned test case. 35 fall scenes were evaluated from a standing position. Out of 35 fall scenes, mail was sent for 31 scenes by the fall detection system. An accuracy of 88.5% was obtained for this test case.

Table 3. Fall events evaluated under broad daylight while standing

Test Cases	Actual Fall	Detected Fall	Accuracy
Total Cases	35	31	88.5

c. Fall events which occurred under low light test case

Table 3 depicts the result for the above-mentioned test case. The main reason for this test case is because falls mainly occur due to poor lighting which distorts the vision. 20 fall scenes were evaluated under low light, out of which only 14 were detected as a fall. The accuracy produced by the system for the above test case is 70%. Accuracy can be improved by using high resolution cameras.

Table 4. Fall events evaluated under low light

Test Cases	Actual Fall	Detected Fall	Accuracy
Total Cases	20	14	70

Accuracy of the System:

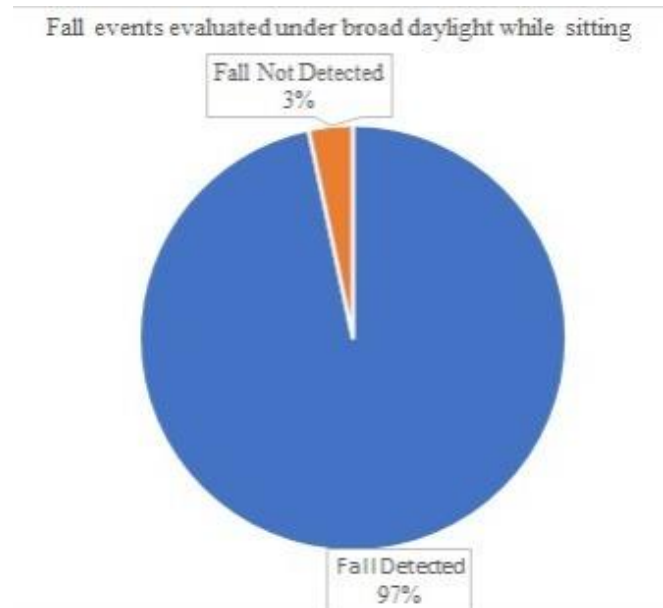


Fig 3. Fall events evaluated under broad daylight while sitting

From Fig 3, it can be seen that the system performs exceptionally well under the test case where the person is sitting and undergoes a fall in broad daylight.

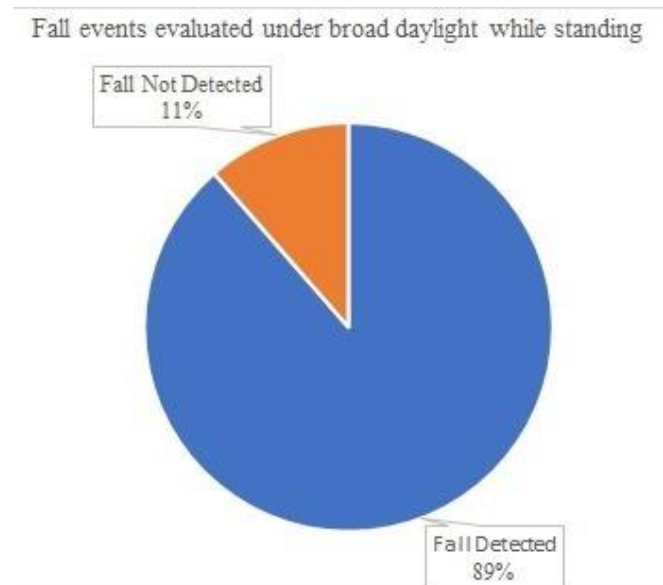


Fig 4. Fall events evaluated under broad daylight while standing

From Fig 4, it can be inferred that the system continues to perform exceptionally well in the case of broad daylight, although the standing test case does not give as great a result as in the case of the sitting test case.

Fall events evaluated under low light

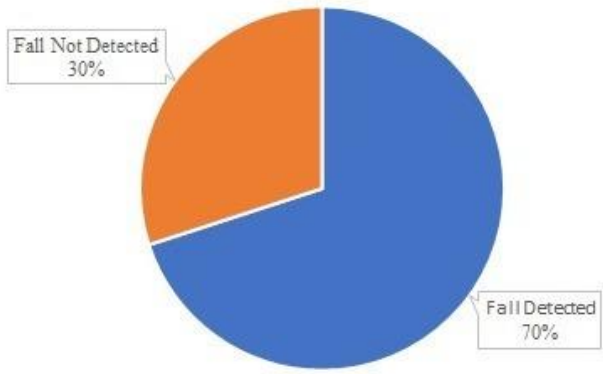


Fig 5. Fall events evaluated under low light

From Fig 5, it can be observed that the performance of the system significantly drops in the case of low lighting in the surroundings. It is still a decent accuracy with a fall detection rate of 70%.

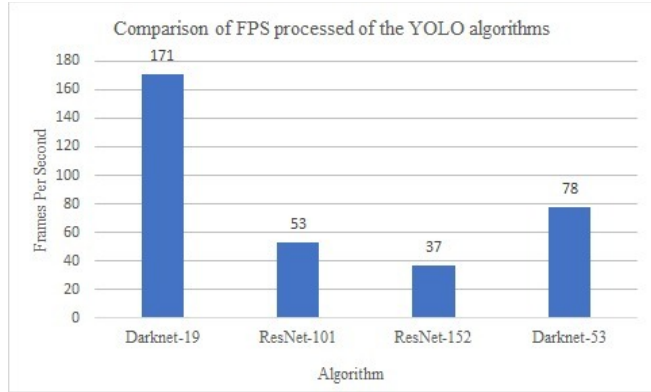


Fig 6. Comparison of FPS processed of the YOLO algorithms

From Fig 6, it can be inferred that the Darknet-19 algorithm gives the highest processing power for performing the YOLO algorithm to detect the person in the frame. The Darknet-19 algorithm gives more than double the speed of processing as compared to any other algorithm.

V. RESULTS

The model detects items in the frame of films or photos in real time. The system analyses each frame of photos and generates residual blocks of various sizes, which it then classifies using the weights it has been trained on. As a result, after the model recognizes our frames, it uses the model to forecast the object in the frame. Because the model is extremely precise, it classifies all of the conceivable objects it can see in the camera, so we set a threshold for it to only detect people in the image. The following step is to determine the distance between the identified object and the camera. We utilize an approximation algorithm to compute the distance using only a single web camera because we don't have access to a depth camera, which is needed for professional depth estimation for self-driving automobiles. This is accomplished by saving reference images of the object's computed distance from the camera, as well as the object's width, into variables for subsequent use.

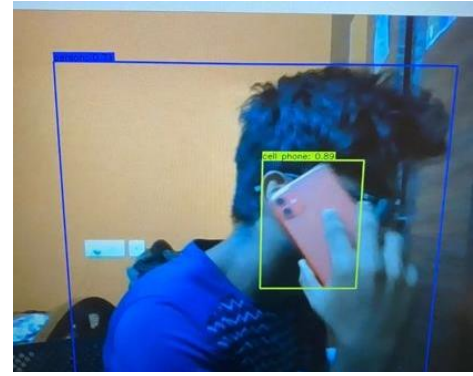

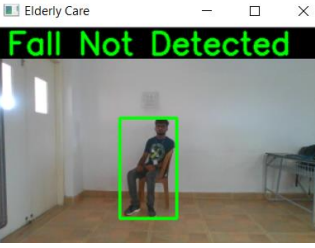


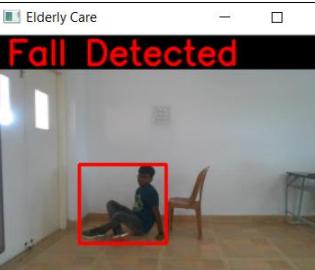



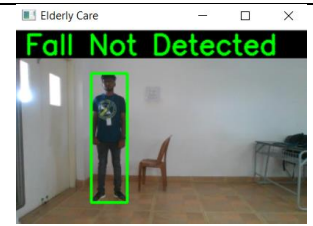
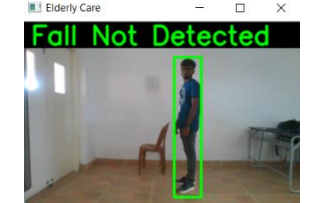

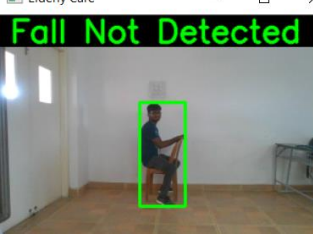



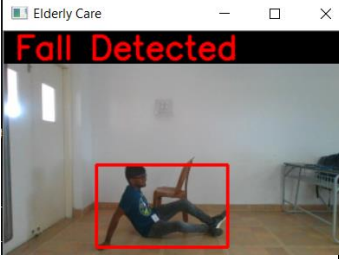
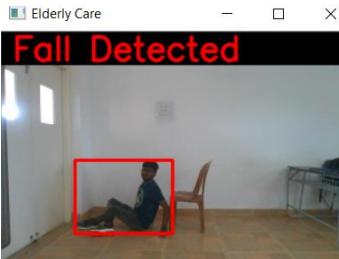
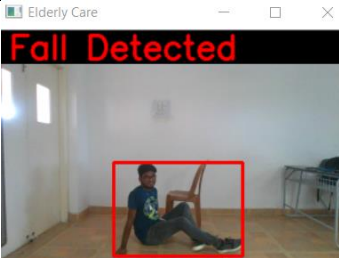

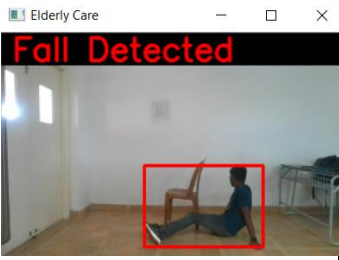
Fig 6. Real time Object Detection scenarios using the proposed object detection model

Table 5. Results of fall-detection Module

Case	Output	Accuracy
Walking (#1)		Correct
Walking (#2)		Correct

Sitting (#3)		Correct
Sitting (#4)		Correct
Sitting (#5)		Correct
Falling (#6)		Correct
Falling (#7)		Correct
Falling (#8)		Correct

Falling (#9)		Correct
Falling (#10)		Correct
Walking (#11)		Correct
Walking (#12)		Correct
Sitting (#13)		Correct
Sitting (#14)		Correct
Sitting (#15)		Correct

Falling (#16)		Correct
Falling (#17)		Correct
Falling (#18)		Correct
Falling (#19)		Correct
Falling (#20)		Correct

VI. CONCLUSIONS

In this work, we have developed fall detection and object detection models using YOLOv4 algorithm and OpenCV to detect the falls in the elderly and the object distance for the blinds through videos and in real-time. Given a video, the model can track a person, position based on the relationship between the person and the floor and finally determines whether the person falls and the distance between him and an object. Our model can be easily integrated into any surveillance systems or webcams to help caregivers and security staff quickly detect and prevent falls. As a result, our

proposed systems can be used at homes, nursing homes to improve the quality of residential care facilities. In the future, this project can be implemented to predict falls instead of detecting the fall after it has happened. Using the emotions of the person, we can predict their heart rate and steadiness of the person to predict the fall before the fact.

REFERENCES

- [1] Foroughi, H., Aski, B.S. and Pourreza, H., 2008, December. Intelligent video surveillance for monitoring fall detection of elderly in home environments. In *2008 11th international conference on computer and information technology* (pp. 219-224). IEEE.
- [2] Wang, J., Zhang, Z., Li, B., Lee, S. and Sherratt, R.S., 2014. An enhanced fall detection system for elderly person monitoring using consumer home networks. *IEEE transactions on consumer electronics*, 60(1), pp.23-29.
- [3] Saleh, M. and Jeannès, R.L.B., 2019. Elderly fall detection using wearable sensors: A low cost highly accurate algorithm. *IEEE Sensors Journal*, 19(8), pp.3156-3164.
- [4] Yacchirema, D., de Puga, J.S., Palau, C. and Esteve, M., 2019. Fall detection system for elderly people using IoT and ensemble machine learning algorithm. *Personal and Ubiquitous Computing*, 23(5), pp.801-817.
- [5] Saleh, M. and Jeannès, R.L.B., 2018, September. An efficient machine learning-based fall detection algorithm using local binary features. In *2018 26th European Signal Processing Conference (EUSIPCO)* (pp. 667-671). IEEE.
- [6] Khraief, C., Benzarti, F. and Amiri, H., 2020. Elderly fall detection based on multi-stream deep convolutional networks. *Multimedia Tools and Applications*, 79(27), pp.19537-19560.
- [7] Zhang, F., Li, Y., You, S. and Fu, Y., 2021. Learning temporal consistency for low light video enhancement from single images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4967-4976).
- [8] Yacchirema, D., de Puga, J.S., Palau, C. and Esteve, M., 2018. Fall detection system for elderly people using IoT and big data. *Procedia computer science*, 130, pp.603-610.
- [9] Badgujar, S. and Pillai, A.S., 2020, July. Fall Detection for Elderly People using Machine Learning. In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-4). IEEE.
- [10] Santiago, J., Cotto, E., Jaimes, L.G. and Vergara-Laurens, I., 2017, January. Fall detection system for the elderly. In *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 1-4). IEEE.
- [11] Sowmyayani, S., Murugan, V. and Kavitha, J., 2021. Fall detection in elderly care system based on group of pictures. *Vietnam Journal of Computer Science*, 8(02), pp.199-214.
- [12] Sumiya, T., Matsubara, Y., Nakano, M. and Sugaya, M., 2015. A mobile robot for fall detection for elderly-care. *Procedia computer science*, 60, pp.870-880.
- [13] Delahoz, Y.S. and Labrador, M.A., 2014. Survey on fall detection and fall prevention using wearable and external sensors. *Sensors*, 14(10), pp.19806-19842.
- [14] Wu, F., Zhao, H., Zhao, Y. and Zhong, H., 2015. Development of a wearable-sensor-based fall detection system. *International journal of telemedicine and applications*, 2015.
- [15] Kwolek, B. and Kepski, M., 2014. Human fall detection on embedded platform using depth maps and wireless accelerometer. *Computer methods and programs in biomedicine*, 117(3), pp.489-501.
- [16] Wadmare, J., Shah, D., Diwakar, C. and Bhanushali, V., 2021. A Vision Based approach for Fall Detection System for Elderly People. Available at SSRN 3867407.
- [17] Nooruddin, S., Islam, M.M. and Sharna, F.A., 2020. An IoT based device-type invariant fall detection system. *Internet of Things*, 9, p.100130.

- [18] Bansal, D., Alsadoon, A., Prasad, P.W.C., Paul, M. and Elchouemi, A., 2018. Elderly people fall detection system using skeleton tracking and recognition. *American Journal of Applied Sciences*, 14(9), pp.423-431.
- [19] Agrawal, S.C., Tripathi, R.K. and Jalal, A.S., 2017, July. Human-fall detection from an indoor video surveillance. In *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-5). IEEE.
- [20] YE, B. and Yu, L. . (2019) "Fall detection system based on BiLSTM neural network", *Transactions on Machine Learning and Artificial Intelligence*, 7(5), pp. 01–12. doi: 10.14738/tmlai.75.6318.
- [21] Khawandi, S., Daya, B. and Chauvet, P., 2010. Automated monitoring system for fall detection in the elderly. *International Journal of Image Processing (IJIP)*, 4(5), p.476.