

Vision-Based Deep Learning Fall Detection System for Elderly People

¹Dr. Harisha S B, ²Anandraj S N, ³Lakshith H B, ⁴Jagadeesh Ganapati Bhat, ⁵Nagaraj C V

¹Assistant Professor, ²Assistant Professor, ³UG Student, ⁴UG Student, ⁵UG Student

¹Electronics and Telecommunication,

¹Jawaharlal Nehru New College Of Engineering, Shivamogga, India

Abstract:

This paper presents a vision-based fall detection system designed to enhance the safety and well-being of elderly individuals through real-time monitoring and rapid emergency response. Falls among the elderly represent a major global health concern, often leading to serious injuries, long-term disabilities, and reduced quality of life, particularly for individuals living alone. Conventional fall detection methods, such as wearable sensors and manual supervision, are limited by user discomfort, high costs, and a high rate of false alarms. To address these challenges, the proposed system integrates the YOLO object detection algorithm with Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to accurately detect fall events from continuous video streams. The video data undergo pre-processing steps including frame extraction, normalization, and background noise reduction. YOLO is employed to detect and localize the human subject in each frame, while the CNN-LSTM architecture learns discriminative spatiotemporal features corresponding to fall-related motion patterns, such as abrupt posture changes, rapid downward movements, and abnormal body orientations. The system is trained and evaluated using annotated fall and non-fall video sequences, achieving high accuracy, sensitivity, and specificity while maintaining real-time performance. Furthermore, the proposed approach is robust to variations in lighting conditions, environments, and subject appearances. As a non-intrusive and cost-effective solution, the system is well suited for deployment in smart homes, intelligent healthcare environments, and assistive living applications. The results demonstrate the potential of the proposed method to enable timely medical intervention and significantly improve the quality of life for elderly individuals.

Key words: Elderly Monitoring, Fall Detection, LSTM, Deep Learning, Vision-Based System, Human Activity Recognition, Spatiotemporal Features, Sequence Modelling, Real-Time Detection, Smart Healthcare, Assistive Technology.

I. INTRODUCTION

Falls among the elderly are one of the most serious global public health concerns because they often lead to severe injuries, disabilities, reduced mobility, and even fatalities. According to the World Health Organization (WHO), nearly one-third of individuals aged 65 years and above experience at least one fall annually, making effective fall detection and prevention a critical component of elder care [1]. Ensuring timely detection is essential to enable quick medical response and prevent complications such as fractures, head trauma, and loss of independence.

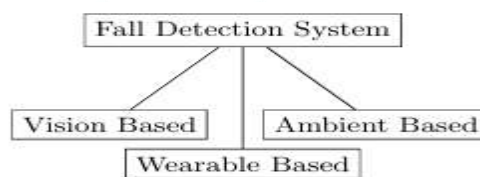


Fig. 1: Various types of a fall detection system

The various fall detection systems are shown in Fig. 1. Wearable sensor-based systems frequently suffer from discomfort, high maintenance, and poor user compliance, as elderly individuals may forget or refuse to wear the device continuously [3, 8]. Ambient-based systems such as infrared or floor vibration sensors provide limited coverage and are costly to deploy. Traditional vision-based systems also struggle to differentiate normal daily actions like sitting or lying down from actual falls due to insufficient temporal motion understanding [2, 4, 5]. These constraints create an urgent need for a robust, non-intrusive, and real-time monitoring solution. Hence, existing fall-detection solutions face notable limitations.

Recent advancements in Artificial Intelligence (AI) and deep learning offer promising solutions to overcome these challenges. In particular, Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have shown exceptional capability in learning complex spatial and temporal features from video sequences [11, 12]. CNNs effectively extract posture and body orientation from individual frames, while LSTMs excel in modelling motion sequences over time.

Combining these architectures enables vision-based fall-detection systems to accurately analyze human movement and reduce false alarms [7, 9, 13, 16]. Vision-based approaches are widely preferred because they are non-intrusive and eliminate dependency on user compliance - a major drawback in wearable systems [8, 15]. Cameras can operate continuously without disrupting the user's comfort, while deep learning models can interpret fall dynamics more reliably than handcrafted feature-based methods [14, 15].



Additionally, visual data offers richer contextual information such as fall speed, posture changes, and final body position, enhancing decision-making accuracy [11].

This paper focuses on designing an intelligent, vision-based fall detection system that uses a web camera for continuous monitoring and a CNN-LSTM model to classify fall and non-fall events in real time. To ensure immediate assistance, the system integrates a GSM module that automatically alerts caregivers via SMS upon detecting a fall, thereby providing a practical and cost-effective safety solution suitable for home and healthcare environments. The proposed system aims to achieve high detection accuracy, reduce false alarms, enhance user comfort, and deliver reliable performance across different activities, lighting conditions, and surroundings.

The rest of the paper is organized as follows: Section II describes the methodology adopted to address the problem; Section III explains the implementation of the proposed system; Section IV presents and discusses the obtained results, finally Section V concludes the paper with key findings and future research directions.

II. METHODOLOGY

The proposed vision-based system for detecting falls in elderly people works in real time and is made up of different parts, as shown in the Fig. 2. A webcam constantly records live video frames and sends them to the Raspberry Pi 5, which is the main controller. The Raspberry Pi 5 gets its power from a regulated power source. The collected frames are processed using deep learning models for person detection, posture analysis, and fall categorization. Upon confirming a fall incident, the Raspberry Pi activates a buzzer for immediate local notice and simultaneously broadcasts emergency messages using a GSM module via a USB-to-UART link, ensuring speedy notification and reaction.

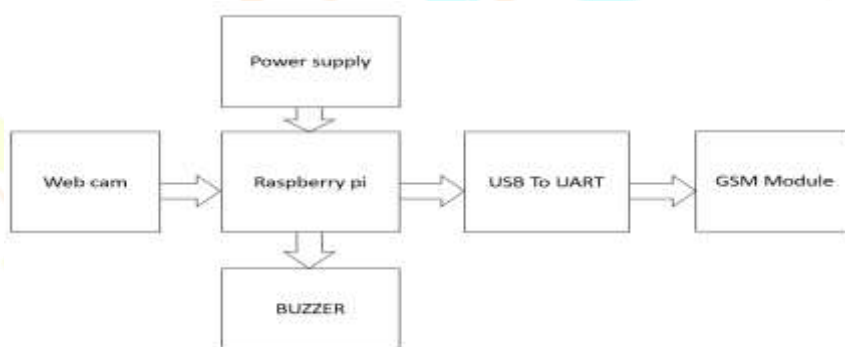


Fig. 2: Block diagram of proposed system

The fall detection algorithm using deep learning is as shown in Algorithm 1, which comprises of following stages:

2.1 Video Data Acquisition

The setup uses a USB webcam attached to the Raspberry Pi to continuously record footage in real time. In order to accurately extract body key points and perform posture analysis, the camera is positioned to capture the old person's entire body. The main input for the fall detection procedure is the recorded live video stream. This vision-based method makes it possible to continuously and non-intrusively monitor senior citizens without the need for wearable sensors.

2.2 Fall Detection Using Body Key Points

The YOLO v3 model receives the pre-processed video frames in order to detect the human subject in real time. Pose estimate is carried out by extracting essential body points, including the head, shoulders, elbows, wrists, hips, knees, and ankles, after the person has been identified. A skeletal model of the human posture is created by combining the spatial coordinates that represent each important point. These normalized key point coordinates are utilized to calculate posture-related characteristics that accurately depict the elderly person's physical posture, such as body tilt, joint angles, and vertical displacement.

2.3 Temporal Motion Analysis Using CNN-LSTM

To learn spatial posture patterns, a CNN is first used to process the normalized body key point information. After being retrieved from successive frames, the spatial information are arranged into temporal sequences and input into the LSTM network. In order to simulate motion patterns, the LSTM examines changes in body key point positions across time. While fall events reveal abrupt posture collapse and fast vertical displacement of key points, normal activities including walking, sitting, and bending show gradual key point transitions.

2.4 Fall Detection and Emergency Alert Generation

The probability value that represents the possibility of a fall occurrence is produced by the LSTM network's output. The system verifies that a fall has occurred if the calculated fall probability is greater than a predetermined threshold value. The registered caregiver or family member receives an emergency SMS alert from the GSM module when a fall is detected. Concurrently, a buzzer attached to the Raspberry Pi is triggered to deliver an instantaneous local audio warning. RealVNC Viewer is used to remotely display the system status and detection results for ongoing monitoring.

Input: Live video stream V
Output: Alert notification if fall is detected

Start

1. Set up the buzzer, GSM module, YOLOv3 model, webcam, and Raspberry Pi.
2. Continue to record live video frames.
3. Prepare frames:
 - a. Adjust the frame size
 - b. Adjust the pixel values
 - c. Keep the FPS constant
4. Use YOLOv3 to identify each frame's human subject.
5. Use the identified person to estimate their pose:
 - a. Identify the head, shoulders, elbows, hips, knees, and ankles;
 - b. Create a representation of skeleton poses.
6. Using successive frames, create a temporal key point feature sequence.
7. Give the trained CNN-LSTM model the key point sequence.
8. Determine p_f , the fall probability.
9. In the event that the expected activity class is "fall,"
 - a. Turn on the GSM module;
 - b. Notify the caregiver via SMS;
 - c. Turn on the buzzer.
 - d. On the monitoring system, display a warning message
10. Else:
 - Continue normal activity monitoring.
11. Repeat steps 3–10 for continuous operation.
12. End

III. IMPLEMENTATION

The implementation of the proposed vision-based elderly fall detection system using an integrated hardware and software architecture is discussed in this section.

3.1 System Architecture Implementation

An integrated hardware–software architecture comprising a Raspberry Pi 5, USB webcam, GSM module, and buzzer is used to create the suggested vision-based fall detection system. The Raspberry Pi uses deep learning models to process the live video frames that the webcam continuously records. While CNN captures spatial posture information, the YOLO method is used for real-time human detection. To distinguish between fall and non-fall activities, the LSTM network examines the temporal motion sequence. When the GSM module detects a fall, the buzzer is turned on for an instant local warning and an emergency alert message is sent to the registered caregiver. Elderly people can be totally automated and monitored in real time thanks to this architecture.

3.2 Hardware Implementation

The system's core processing unit, the Raspberry Pi 5, is in charge of communication, sensor control, and deep learning inference. The Raspberry Pi is interfaced with a high-resolution USB webcam to record continuous video input. In order to transmit emergency SMS alerts during fall incidents, the GSM module is connected via serial connection. GPIO pins are used to interface a buzzer in order to provide a locally audible alarm. A regulated power source powers every component, guaranteeing steady operation for ongoing real-time monitoring.

Our main computing platform is a Raspberry Pi 5, which is powered by a 2.4 GHz quad-core 64-bit ARM Cortex-A76 processor, a Video Core VII GPU, LPDDR4X RAM, and quick microSD storage. A USB camera (a 12-megapixel model from Lapcare) connected to one of the Pi's USB ports continually captures video frames for real-time fall detection using a deep learning model. A SIM800L GSM module with an antenna and an additional Pi USB port connected by a USB-to-UART connection are also included in the system. This connection strategy allows serial communication for SMS-based emergency warnings by connecting the TX pin of the GSM module to the UART RX of the GSM RX to UART TX with a common ground and common VCC. When a fall is detected, a buzzer is linked to GPIO pin 40 to provide a local aural alert. A 5 V power bank powers the complete arrangement, guaranteeing portability and continuous operation without the need for a fixed power source.

3.3 Software Implementation

The whole software implementation of the suggested vision-based fall detection system, including the software tool, algorithm, mathematical formulation, and real-time decision logic, is covered in this part. Python is used to construct the system, which combines embedded hardware control, deep learning, and computer vision.

The following libraries and tools are used in the implementation of the suggested system.

Software tools are used in the suggested system.

- OpenCV – Video capture, preprocessing, and visualization
- YOLOv3 – Real-time human detection
- MediaPipe Pose – Skeletal keypoint extraction
- TensorFlow/Keras – CNN-based activity classification
- GPIO Library – Buzzer control

Mathematical Formulation

The following mathematical equations are used during software execution:

$$1. \text{ Frame Normalization } x_n = \frac{x}{255}$$

Where x is the original pixel value and x_n is the normalized pixel value.

$$2. \text{ Frame Per Second (FPS) } FPS = \frac{1}{T}$$

Where T is the processing time for one frame.

$$3. \text{ YOLO Bounding Box Computation : Center}_x = d_0 \times W, \text{ Center}_y = d_1 \times H$$

$$w = d_2 \times W, h = d_3 \times H$$

$$x = \text{Center}_x - \frac{w}{2}, y = \text{Center}_y - \frac{h}{2}$$

Where W and H are frame width and height.

$$4. \text{ Key point Normalization } k_n = \frac{k}{\max(k)}$$

$$5. \text{ Activity Classification Using Arg Max Predicted Class} = \arg \max(P_i)$$

Where P_i is the output probability of the CNN classifier.

$$6. \text{ Joint Angle Calculation } \theta = \left| \tan^{-1}\left(\frac{y_c - y_b}{x_c - x_b}\right) - \tan^{-1}\left(\frac{y_a - y_b}{x_a - x_b}\right) \right| \times \frac{180}{\pi}$$

$$\text{If } \theta > 180^\circ \Rightarrow \theta = 360^\circ - \theta$$

3.4 YOLO-Based Person Detection Module

The YOLO v3 (You Only Look Once version 3) algorithm is used to identify people in live webcam video broadcasts. The YOLO network receives each video frame, recognizes the human subject, and creates bounding boxes around those individuals. Only pertinent human regions are processed further for posture analysis thanks to this phase. YOLO is appropriate for real-time execution on the Raspberry Pi platform due to its high detection speed and precision. The more details about the YOLO architecture in Fig. 3 can be found in [17].

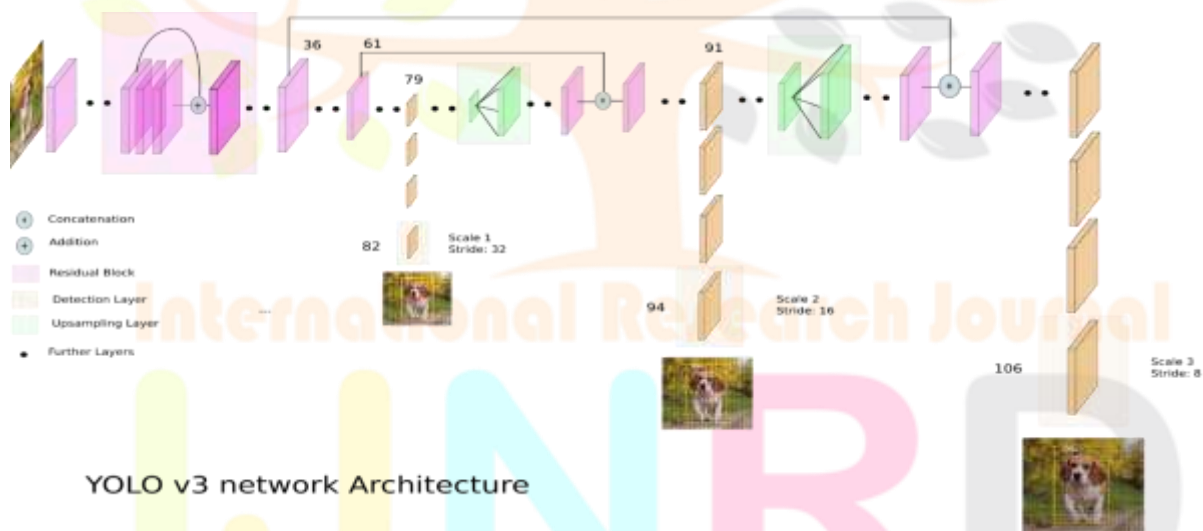


Fig. 3: YOLO v3 Architecture [17]

Key Features of YOLOv3

1. **Architecture:** The core of YOLO v3 is a deep convolutional neural network known as Darknet-53. This network is strong and effective for feature extraction since it comprises 53 convolutional layers with skip connections, just like Res Net.
2. **Multi-Scale Detection:** YOLO v3 can identify objects of varied sizes by predicting bounding boxes at three different scales. A feature pyramid network (FPN) technique is used to accomplish this.
3. **Bounding Box Predictions:** Every grid cell forecasts several bounding boxes, each with coordinates, object confidence, and class probability. To get rid of unnecessary detections, Non-Maximum Suppression (NMS) is used.
4. **Accuracy and Speed:** YOLO v3 strikes a balance between the two. It maintains a high mean average precision (mAP) of 87.54% on the VOC dataset while processing images at 30 frames per second on a Titan X GPU.
5. **Loss Function:** YOLO v3 employs a unique loss function that incorporates terms for classification, object confidence, and bounding box regression. This guarantees accurate classification and localization.
6. **Flexibility:** YOLO v3 is appropriate for datasets with overlapping labels since it offers multi-label categorization. Instead of using SoftMax for class predictions, it employs binary cross-entropy loss, enabling more complex classifications.

3.5 CNN-Based Posture Feature Extraction



The cropped human region is sent to the Convolutional Neural Network (CNN) for spatial feature extraction following human detection. By learning important spatial patterns including body orientation, joint positions, and postural alignment, the CNN assesses body posture. The retrieved features provide a succinct numerical representation of the individual's physical position. The LSTM network receives these information and uses them to analyze temporal motion. By removing background noise and concentrating just on human position, CNN increases classification accuracy.

3.6 LSTM-Based Temporal Motion Analysis

Sequential posture features are analysed over time by the Long Short-Term Memory (LSTM) network. LSTM efficiently records temporal relationships across consecutive frames because fall detection relies on abrupt motion changes. The LSTM model receives the CNN-extracted posture features from several frames organized as time sequences. The motion is then categorized by the LSTM as either a fall or a typical action like sitting, bending, or walking.

3.7 Fall Detection Decision Logic

The suggested implementation employs the Arg Max technique as a direct class-based choice mechanism. Without using a probability threshold, the system instantly initiates the warning mechanism if the CNN classifier predicts the activity label as "fall".

IV RESULT AND DISCUSSION

The experimental findings and performance assessment of the suggested vision-based fall detection system are presented in this part.

4.1 Experimental Setup

A Raspberry Pi 5, USB webcam, GSM module, and buzzer were used in an interior residential setting to conduct an experimental evaluation of the suggested vision-based geriatric fall detection system. In order to record the subject's entire body motions, the camera was positioned at an elevated angle. Under regular lighting circumstances, a variety of test scenarios were carried out, including walking, sitting, bending, lying down, and unexpected falls. Real VNC Viewer, which was installed on a personal computer, was used to remotely execute and monitor the live system.

4.2 Dataset and Activity Scenarios

Both real-time recorded video clips and sample fall activity datasets from the internet were used to evaluate the system. Walking, sitting in a chair, lying on a bed, bending to select objects, and unintentional forward and backward falls were all included in the activity set. These exercises were chosen to accurately depict the movements of senior citizens. Classification was assessed using both fall and non-fall activities.



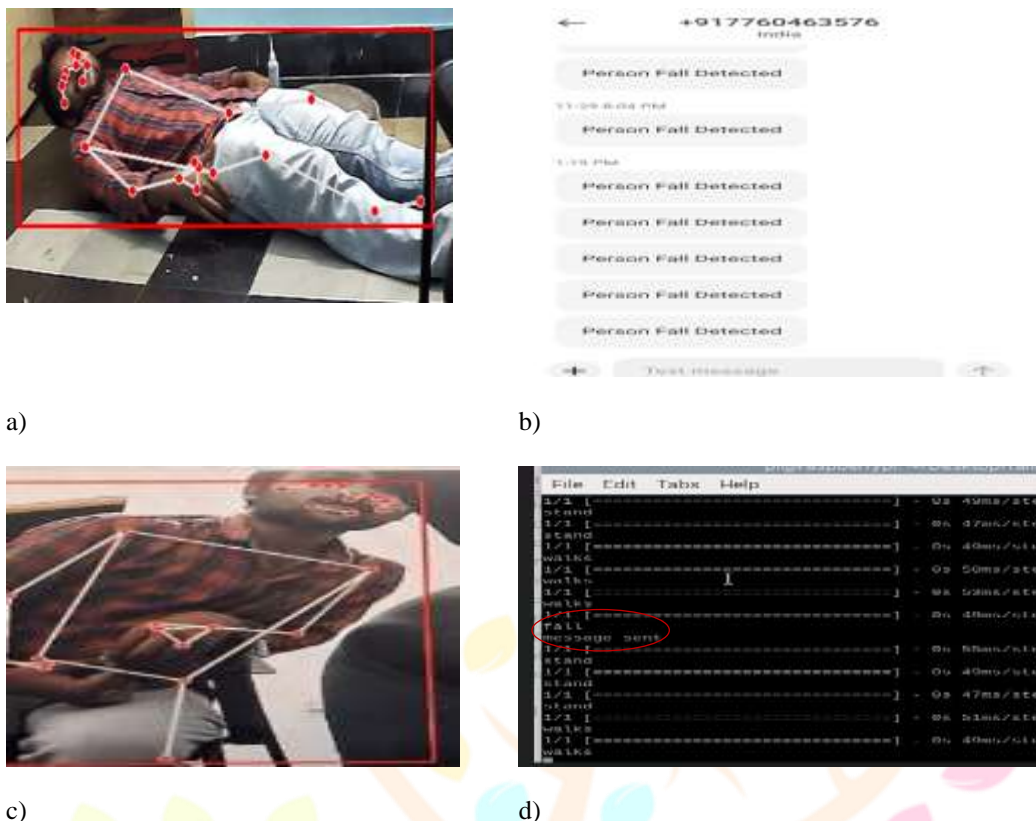


Fig. 4: Results of the proposed method. a) fall test image b) message sent to the caregiver c) non fall test image - d) Real VNC viewer snapshot confirming the fall .

The proposed approach properly identified the image in Fig. 4a as a fall occurrence. The human body in a completely horizontal position on the ground was precisely localized using the YOLO algorithm. Critical posture features that indicate full body collapse and ground-level alignment were extracted by CNN. The proposed method identified an event as a fall because the LSTM model identified a quick temporal shift in body orientation and vertical displacement that closely matched the fall motion patterns found in the training dataset. Fig. 4b shows the notification provided to the caregiver following the fall detection.

The system properly identified the real-time image in Fig. 4c as a seated posture. The human subject was successfully identified by the YOLO model, which also produced a precise bounding box around the body. Posture characteristics including bent knees, forward torso inclination, and stable body orientation were collected by CNN. After analysing the temporal motion, the LSTM verified that there was no abrupt vertical displacement or fast postural collapse. The action was appropriately categorized as a non-fall (sit) event based on these spatial-temporal characteristics.

The system continuously analyzes human posture states such as stand and walks frame by frame, with each state processed when an abnormal posture sequence corresponding to a fall is detected, the system successfully triggers the alert mechanism. This event is confirmed in the log by the message “message sent” as highlighted in Fig. 4d

The efficacy of the suggested YOLO–CNN–LSTM-based system is demonstrated by the accurate classification of the sitting posture as a non-fall case and the laying posture as a confirmed fall case. The outcomes confirm that the model does not mix up everyday tasks like sitting with real fall incidents. The trained system operates reliably under realistic indoor settings, as demonstrated by the successful real-time prediction seen using Real VNC Viewer. These findings confirm the system's dependability and suitability for monitoring senior falls in homes in real time.

4.3 Inter mediate results of YOLO,CNN and LSTM models

The YOLOv3 model's function in the fall detection system is to identify people in real time using live video frames that the webcam captures. By constructing a bounding box around the human subject and disregarding the background in Fig 5, it precisely locates and identifies the subject. This guarantees that only pertinent human areas are sent on to the subsequent processing phase. The accuracy and speed of the entire fall detection system are increased by YOLO's continuous tracking of the subject in each frame, which gives CNN-based posture feature extraction and LSTM-based motion analysis a clear and focused input.



Fig. 5: Image showing Human detection

The pose estimation module receives the identified human region to extract skeletal key points once the individual has been localized using YOLOv3. Major body joints including the head, shoulders, elbows, hips, knees, and ankles are included in the extracted position. Before being fed into the CNN classifier, these important points are normalized and reshaped. The CNN picks up spatial posture characteristics such as limb placement, joint alignment, and body orientation. The skeletal pose overlay effectively depicts the human subject's postural structure, as seen in Fig. 6. Based on spatial body configuration, this intermediate output aids CNN in differentiating between actions including sitting, walking, standing, and falling.

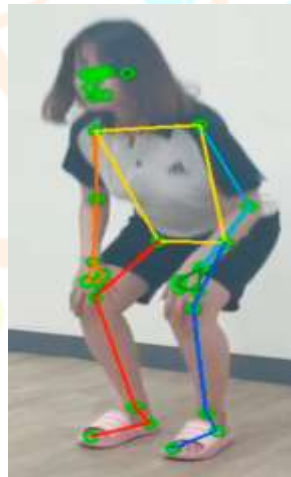


Fig. 6: Image showing key points extraction skeletal by CNN

The LSTM network is used for temporal motion analysis in the last intermediate step. The LSTM model receives the sequential key points features that were taken from successive frames and organized into a temporal sequence. This makes it possible for the system to examine motion dynamics, including abrupt posture changes, body movement direction, and fall speed. The probability distribution across predetermined activity classes is represented by the LSTM output. The shift from upright to ground-level posture, a crucial sign of a fall event, is well captured by the temporal skeleton motion sequence, as seen in Fig. 7.

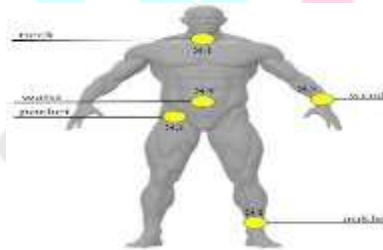


Fig. 7: Reference Body Points for Temporal Movement Analysis in Fall Detection Systems

4.4 Real-Time Monitoring Using Real VNC Viewer

Real VNC Viewer was used to successfully accomplish remote system monitoring and real-time visualization. A remote PC was used to continuously monitor the live video feed, detection bounding boxes, fall classification outputs, and system status. This made it possible for the operator to keep an eye on the old person's health without having to be close to the Raspberry Pi device. In home healthcare settings, Real VNC Viewer's remote access feature improved system usability, security, and implementation simplicity.

4.5 GSM Alert and Buzzer Response Analysis

The GSM module successfully sent emergency alert messages to the registered mobile number in a matter of seconds after confirming fall detection. The alarm message gave caregivers instant notice so they could act quickly. The buzzer simultaneously produced a

loud local alarm for help in the vicinity. The system's overall dependability was increased by the dual alert mechanism, which guaranteed both distant and prompt local emergency reaction.

V CONCLUSION

This paper presented a vision-based fall detection system aimed at improving the safety, independence, and quality of life of elderly individuals through continuous and real-time monitoring. By integrating YOLO for efficient human detection with a CNN-LSTM architecture for spatiotemporal feature learning, the system successfully addresses the limitations of traditional fall detection approaches such as wearable sensors and manual supervision. The proposed method accurately captures critical fall-related characteristics, including sudden posture transitions, rapid downward movements, and abnormal body orientations, enabling reliable differentiation between fall and non-fall activities. Comprehensive preprocessing techniques, including frame extraction, normalization, and background noise reduction, further enhance detection accuracy and robustness. Experimental evaluation on annotated video datasets demonstrates high accuracy, sensitivity, and specificity while maintaining real-time performance, making the system suitable for practical deployment. Additionally, the non-intrusive nature of the vision-based approach ensures user comfort and compliance, especially for elderly individuals living alone. The system's robustness to variations in lighting conditions, environments, and subject appearances highlights its adaptability to real-world scenarios. Overall, the proposed fall detection framework represents an effective and scalable solution for smart homes and intelligent healthcare environments, with significant potential to reduce response time during emergencies, prevent severe injuries, and support independent living among the elderly population. Future enhancements may include integration with IoT and cloud platforms for remote monitoring and automated alerts to caregivers. Expanding the system to support multi-person detection, adaptive learning models, and privacy-preserving techniques can further improve accuracy, scalability, and acceptance in real-world healthcare applications.

REFERENCE

1. El-Bendary, Nashwa, Qing Tan, Frédérique C. Pivot, and Anthony Lam. "Fall detection and prevention for the elderly: A review of trends and challenges." *International Journal on Smart Sensing and Intelligent Systems* 6, no. 3 (2013): 1230.
2. Anishchenko, L., Zhuravlev, A., Chizh, M.: Fall detection using multiple bioradars and convolutional neural networks. *Sensors* 19(24), 5569 (2019)
3. Kaewkannate, K., Kim, S.: A comparison of wearable fitness devices. *BMC public health* 16, 1–16 (2016)
4. Kwolek, B., Kepski, M.: Fall detection using kinect sensor and fall energy image. In: *Hybrid Artificial Intelligent Systems: 8th International Conference, HAIS 2013, Salamanca, Spain, September 11-13, 2013. Proceedings* 8. pp. 294–303. Springer (2013)
5. Sase, P.S., Bhandari, S.H.: Human fall detection using depth videos. In: *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*. pp. 546–549. IEEE (2018)
6. Taramasco, C., Rodenas, T., Martinez, F., Fuentes, P., Munoz, R., Olivares, R., De Albuquerque, V.H.C., Demongeot, J.: A novel monitoring system for fall detection in older people. *Ieee Access* 6, 43563–43574 (2018)
7. F. Castro, V. Dentamaro, V. Gattulli, and D. Impedovo, "Fall Detection with LSTM and Attention Mechanism," *CEUR Workshop Proceedings*, vol. 3517, pp. 1-10, 2023.
8. D. Kumar, A. K. Ravikumar, V. Dharmalingam, and V. P. Kafle, "Elderly Health Monitoring System with Fall Detection Using Multi-Feature Based Person Tracking," *Proceedings of the 2019 International Conference on Intelligent Computing and Control Systems*, 2019.
9. G. Anitha and S. B. Priya, "Vision Based Real Time Monitoring System for Elderly Fall Event Detection Using Deep Learning," *Computer Systems Science & Engineering*, vol. 37, no. 3, pp. 317-331, 2022.
10. D.-W. Lee, K. Jun, K. Naheem, and M. S. Kim, "Deep Neural Network-Based Double-Check Method for Fall Detection Using IMU-L Sensor and RGB Camera Data," *IEEE Access*, vol. 9, pp. 36720-36730, 2021.
11. J. Chen, J. Wang, Q. Yuan, and Z. Yang, "CNN-LSTM Model for Recognizing Video-Recorded Actions Performed in a Traditional Chinese Exercise," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 11, pp. 888-897, 2023.
12. K. Al-Majdi, R. S. H. AL-Musawi, A. H. Ali, and Y. S. Mezaal, "Real-time classification of various types of falls and activities of daily livings based on CNN LSTM network," *Periodicals of Engineering and Natural Sciences*, vol. 9, no. 3, pp. 958-969, 2021.
13. R. Espinosa, H. Ponce, S. Gutiérrez, L. Martínez-Villaseñor, J. Brieva, and E. Moya-Albor, "A vision-based approach for fall detection using multiple cameras and convolutional neural networks: A case study using the UP-Fall detection dataset," *Computers in Biology and Medicine*, vol. 116, 103520, 2020.

14. M. M. Islam, O. Tayan, M. R. Islam, M. S. Islam, S. Nooruddin, M. N. Kabir, and M. R. Islam, "Deep Learning Based Systems Developed for Fall Detection: A Review," IEEE Access, vol. 8, pp. 166113-166137, 2020.
15. S. Hashemi, "Video-based human fall detection in indoor spaces for health monitoring," Ph.D. dissertation, University of British Columbia, Vancouver, 2023.
16. S. Sahni, S. Jain, and S. Khetwat Saritha, "A novel ensemble model for fall detection: leveraging CNN and BiLSTM with channel and temporal attention," Automatika, vol. 66, no. 2, pp. 103-116, 2025.
17. <https://share.google/IjmjRCq7N1K1WfXAM>

