FMDB *Transactions on Sustainable Biomedical Engineering*

Proactive Vision-Based Fall Risk Detection with Human-in-the-Loop Audio Confirmation for Elderly Safety

Chettiyar Vani Vivekanand[1,*], M. Pavithra[2], S. Kalaivani[3], Ciyamala Kushbu[4], E. Anna Devi[5], D. Samson[6]

[1,2]Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, Tamil Nadu, India.
[3]Department of Electronics and Communication Engineering, Rajalakshmi Institute of Technology, Chennai, Tamil Nadu, India.
[4]Department of Information Science and Engineering, CMR Institute of Technology, Bengaluru, Karnataka, India.
[5]Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India.
[6]School of Computer Science and Information Technology, University of Edenberg, Lusaka, Zambia.
chettiyarvanivivekanand@rajalakshmi.edu.in[1], pavithra.m.2025.mecse@rajalakshmi.edu.in[2], kalaivani.s@ritchennai.edu.in[3], ciyamala91@gmail.com[4], annadevimit@gmail.com[5], samson@edenberg.org[6]

Abstract: Since falls are a leading cause of harm and diminished mobility in the elderly population, there is a need for discreet and dependable indoor monitoring systems. Current vision-based methods primarily focus on post-fall detection and often exhibit poor adaptability to individual movement patterns and high false alarm rates. This paper suggests a vision-based, real-time, non-wearable fall risk prediction system that incorporates a human-in-the-loop confirmation mechanism to improve practical usability and proactively estimates fall risk. The system continuously analyses human posture using a single monocular RGB camera and MediaPipe-based skeletal pose estimation. To account for inter-subject variability, an online, personalized baseline posture model is learned without prior calibration. A continuous scoring method based on torso angle deviation and temporal instability features is used to quantify fall risk. Before alert generation, a temporal instability observation window is used to minimize false alarms. Additionally, users can turn off pointless alerts with an audio prompt and gesture-based confirmation that uses right-hand raise detection. The suggested system is suitable for proactive elderly safety monitoring, as evidenced by experimental results from real-time indoor video streams demonstrating reliable fall risk tracking, fewer false alarms, and real-time feasibility.

Keywords: Fall Risk Prediction; Elderly Safety Monitoring; Vision-Based System; Human Pose Estimation; Real-Time Monitoring; Temporal Instability Analysis; Gesture-Based Confirmation.

Received on: 13/05/2025, **Revised on:** 24/07/2025, **Accepted on:** 03/09/2025, **Published on:** 09/03/2026

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSBE

DOI: https://doi.org/10.69888/FTSBE.2026.000669

Cite as: C. V. Vivekanand, M. Pavithra, S. Kalaivani, C. Kushbu, E. A. Devi, and D. Samson, "Proactive Vision-Based Fall Risk Detection with Human-in-the-Loop Audio Confirmation for Elderly Safety," *FMDB Transactions on Sustainable Biomedical Engineering*, vol. 1, no. 1, pp. 1–17, 2026.

1. Introduction

*Corresponding author.

Falls are among the most serious health risks faced by older adults and represent a leading cause of injury, hospitalization, and loss of independent living worldwide. With the rapid growth of the ageing population, falls have become a major public health concern, imposing substantial physical, psychological, and economic burdens on individuals, caregivers, and healthcare systems [1]. To improve elderly safety and the quality of daily life in the living environment, early identification of fall risk and timely intervention are necessary. Traditional fall-monitoring solutions mostly relied on wearable sensors, such as accelerometers and gyroscopes, to detect abrupt movement patterns associated with falls [2]. Yet wearable systems achieve high detection accuracy under controlled conditions; limited user compliance, discomfort, battery maintenance requirements, and improper device placement often compromise their effectiveness in real-world settings. But these limitations are a major drawback for elderly users, who may forget or refuse to wear such sensing devices throughout the day. Thus, a practical alternative is the use of non-wearable monitoring approaches, which have received increasing attention in recent years. Recent advances in computer vision and pose estimation have enabled vision-based fall detection systems, offering a promising, non-intrusive alternative for fall monitoring [3]. Existing vision-based methods can be broadly categorized into post-fall detection and learning-based classification approaches. Post-fall detection systems identify fall events after they occur by analyzing sudden posture changes or detecting ground contact.

This is neither suitable for emergency response nor for preventive measures, as it provides no early warning. Learning-based approaches, particularly deep neural networks, have been proposed for recognizing falls from video data [4]. However, these methods often require large, annotated datasets, are limited in interpretability, and may degrade in performance when deployed in unseen environments or across diverse subject populations. Furthermore, their black-box nature raises concern about reliability and real-time decision-making in safety-critical applications. False alarms caused by voluntary movements, such as bending, sitting, or stretching, that mimic fall motion patterns are an additional challenge for vision-based fall monitoring methods. Often, false alarms question system reliability and lead to alarm fatigue, causing users and even caregivers to ignore or turn off the system entirely. Despite this issue, most existing studies focus primarily on detection accuracy rather than on human-centred interaction and confirmation in practical scenarios. To overcome these limitations, there is an increasing demand for fall-monitoring systems that are proactive rather than reactive, interpretable rather than opaque, and human-centred rather than fully autonomous [5]. Especially proactive fall risk prediction, which seeks to identify instability or hazardous postural dynamics before any fall occurs, provides considerable potential for preventive intervention. At the same time, incorporating human-in-the-loop mechanisms can enable users to validate or cancel system decisions, thereby improving trust and reducing unnecessary alerts. A fall is usually the result of many things that happen before it, like losing balance or having trouble standing up straight. These things can be hard to notice. If researchers notice them, they can take steps to prevent the fall. Systems that can constantly watch people and detect these changes.

The system can find patterns that mean someone might fall. This way, the system can act before the fall occurs or immediately after. This is a lot better than waiting for fall to happen and then doing something about it. Another thing to consider is how to use these systems in real life. For people who live alone and need systems that work without bothering them. Researchers do not want systems that require people to do things all the time or that need to be fixed frequently. Camera-based systems are effective because they can monitor people without requiring them to wear anything. These systems must be good at understanding what people are doing. Sometimes people do things that might look like they are falling. They are not. It is also important for these systems to be able to change and learn. People are all different. Researchers move in different ways. The simplest system model may produce an error or not be suitable for the working environment. The system model should learn about each person over time for improvement. This led to the model being more effective and to fewer mistakes. The people or users have a clear understanding of the system model, and the learning procedure works. If the model has a high decision-making capacity and can handle everything on its own, the user may lead to the most real issues in the system. Suppose users or others can confirm their illness, which might enhance the system model. It may help reduce the false alarm. The system may issue a false alarm, leading people to think the alarm is false, which supports the idea that the system is not smarter; in this way, the system model should work well. Researchers do not want systems that make many false alarms. If this happens, people might start to ignore the alarms. That can be very bad. So, researchers need systems that're smart and can ask people for help.

This way, researchers can be sure that the system is working well and keeping people safe. In the end, researchers need systems that're good at observing people as they learn about them and asking for help when needed. If researchers can make systems like this, they can help older people live better lives. These systems can help prevent falls, reduce hospital visits, and make life easier for everyone. The main thing is that fall monitoring systems need to be designed with people in mind. Researchers need to consider how people live and how they can help them. By doing this, researchers can create systems that work for everyone and really work. Fall monitoring systems like this can make a difference in people's lives. They can help people stay safe and healthy. That is the most important thing. In this work, researchers propose a real-time, non-wearable, vision-based fall-risk prediction framework specifically designed for indoor elderly safety monitoring. The proposed system employs a single monocular RGB camera and uses skeletal pose estimation to analyse human posture dynamics continuously. Unlike prior methods that rely on fixed thresholds or population-level models, the system learns a personalized baseline posture model online, enabling adaptation to inter-subject variability without prior calibration. Fall risk is continuously quantified using torso angle deviation and temporal instability features, allowing early detection of pre-fall conditions. To further enhance reliability, a temporal instability observation window is introduced to suppress transient disturbances. In addition, a human-in-the-loop confirmation mechanism combining audio prompts and gesture-based feedback is integrated to reduce false alarms before alert escalation.

2. Related Work

Research on fall monitoring and detection has received considerable attention in recent years due to the increasing demand for elderly safety solutions. Different approaches have been proposed in the literature to detect falls. In general, these approaches can be grouped into three categories: wearable sensor–based methods, vision-based fall detection systems, and techniques that assess fall risk or predict a fall before it occurs. More recently, attention has also been given to human-centred interaction mechanisms, as researchers seek to make these systems more reliable in real-world environments. Wearable sensor–based methods are used to detect fall events. These systems usually rely on inertial measurement units, such as accelerometers and gyroscopes, that capture sudden changes in body motion. Motion patterns associated with falls are identified by analyzing sensor data. Under controlled experimental conditions, wearable systems have often shown high detection accuracy. This is typically achieved by examining factors such as acceleration magnitude, impact peaks, and variations in body orientation [6]; [7]. In some studies, physiological information is also included to improve the robustness of the detection process. Wearable approaches provide several advantages, their use in real elderly care environments is not always straightforward. One important issue is user compliance. It is not always possible to ensure that the device will be worn continuously, particularly by older adults who may experience cognitive impairment or limited mobility. For example, the device may cause discomfort, the battery needs to be recharged regularly, or the device may be misplaced.

After prolonged device use, the system's reliability can degrade [8]. The best alternative to wearable devices features a vision-based fall-detection system. The cameras were used to detect a person's motion and posture, allowing the fall event to be tracked without any physical contact devices. The vision-based fall detection method relied on computer vision techniques that detect sudden changes in body motion orientation during fall events [9]. These approaches were sensitive to lighting variations, occlusions, and complex indoor scenes. More recent studies leverage advances in human pose estimation to extract skeletal representations and analyze joint trajectories for fall detection [10]. Skeleton-based methods improve robustness by focusing on human posture rather than raw pixel information. In video-based fall detection, deep learning-based approaches, including convolutional and recurrent neural networks, have been widely explored [11]. Although these models may achieve the best performance on the largest labeled dataset, they exhibit limited interpretability and often struggle to generalize across environments, camera views, and other factors. A common challenge in vision-based systems is the post-fall detection model, which only identifies the fall; these features don't provide any advanced warning system to support emergency response in post-fall detection. Preventing intervention may lead to the critical need to reduce the severity of the injury. To overcome the limitations of post-fall detection models, several researchers have investigated pre-fall detection and fall risk assessment models.

These may support deeper analysis by identifying instability, imbalance, or hazardous posture in a dynamic environment before a fall occurs. Some work may analyze patterns using centre-of-mass displacement or postural swaps and may use a deeper sensor or motion-capture model [12]; [13]. These systems may rely entirely on specialized hardware or controlled experimental setups to overcome the limitations of scalability and development in a typical residential environment. The monocular cameras were used for vision-based pre-fall valuation, an area that remains relatively underexplored. Existing models have fixed thresholds or fail to account for inter-subject changeability in movement and posture [14]. The majority focus on offline analysis rather than real-time operation, reducing their applicability to continuous monitoring scenarios. False alarms are a significant problem for both wearable and vision-based fall monitoring systems. Frequent false alarms may result from everyday actions such as sitting, bending, or stretching, which can create motion patterns that mimic falls [15]. Overly frequent false alerts erode user confidence and can lead to alarm fatigue, which ultimately restricts system adoption. A small number of studies have investigated user or human-in-the-loop confirmation as a solution to this problem. After a fall is detected, some systems allow users to turn off alarms by voice commands or manual input [16]. Nevertheless, rather than being incorporated into the fundamental detection logic, these processes are frequently regarded as supplementary elements.

Additionally, rather than using user confirmation during proactive risk assessment, most current systems use it only after post-fall detection. It is evident from examining most current fall detection systems that they were primarily designed to respond after a fall had already occurred. Although this strategy has proven helpful in emergencies, it hasn't done much to avoid injuries beforehand [11]. It has been much more helpful in real-life situations, particularly for older people, if a system can identify early warning indicators, such as a small imbalance or a dangerous posture, before a fall occurs. Another limitation observed is that many systems have treated all users similarly. Vision-based approaches have often relied on fixed thresholds or generalized models. However, human movement varies from person to person. Factors such as age, physical condition, and mobility have influenced how individuals move. Because of this, systems that have not adapted to individual differences have sometimes produced inconsistent or less reliable results. It has also been noticed that many of these models have been evaluated in controlled environments. While such settings have helped in measuring performance, they have not always reflected real-world conditions [15]. In Daily life, environments have clearly shown that changing lighting, occlusions, and unpredictable movements result in systems that perform optimally under certain conditions, not always. Delivering real-time, continuous monitoring is the most important phase of deep learning, which has been widely explored and has achieved strong performance in many cases, despite being criticized for lacking interpretability. In many situations, it has been difficult to explain how a decision has been made clearly. This lack of transparency has eroded trust, particularly in safety-critical applications where learning system behavior has been important.

Researchers have developed a methodology that combines conventional approaches with machine learning techniques to address these issues. For instance, lightweight models have been combined with statistical feature-based techniques to increase efficiency without sacrificing interpretability [17]. In the same way, hybrid frameworks that integrate learning models with rule-based reasoning have improved decision-making. Additionally, multimodal systems—which incorporate many data types such as vision, wearable sensors, and ambient inputs—have been studied [18]. Although these methods have typically improved robustness, they have also increased system complexity and cost, making them less feasible for deployment in typical homes. Edge computing has been receiving more attention lately. These methods involve processing data locally rather than sending it to cloud servers [19]. This has helped reduce latency and improve privacy while balancing computational efficiency, but fall detection accuracy remains a challenge. In addition to relying solely on single-frame notes, temporal modeling techniques have been investigated to analyze movement variations over time. Gradual instability patterns, which frequently precede falls, have been captured using techniques such as recurrent neural networks [20]. Since falls have rarely happened instantly, this strategy has been regarded as more realistic. False alarms remain a major problem despite these developments. Sitting, bending, and stretching are everyday activities that often trigger patterns such as falls, leading to false alarms. Alarm fatigue and a decline in user trust have resulted from this. To address this, several systems have added user confirmation features that let users turn off warnings as needed [21]. These features have often been treated as secondary rather than being fully integrated into the core system. Privacy concerns have also been increasingly important, particularly in camera-based monitoring systems. Continuous video recording has raised concerns about personal space and data security.

To address this, some approaches have used skeletal representations instead of raw video data, helping preserve privacy while maintaining functionality [22]. Overall, although significant progress has been made, several challenges remain. Future systems need to focus on adaptability, real-world reliability, reduced false alarms, and improved privacy protection. Addressing these aspects has been essential to developing practical, widely accepted fall-monitoring solutions. Based on the above review, several key gaps can be identified [23]. First, most existing vision-based systems focus on post-fall detection and lack proactive fall risk prediction capabilities. Second, personalization and online adaptation to individual postural characteristics are rarely addressed, despite significant inter-subject variability among elderly users [24]. Third, false alarm reduction is often treated as a secondary objective, with limited integration of human-centered confirmation mechanisms into real-time decision-making. In contrast, the proposed work aims to bridge these gaps by introducing a real-time, monocular vision-based fall risk prediction framework that combines personalized posture modeling, temporal instability analysis, and a human-in-the-loop audio–gesture confirmation mechanism. By emphasizing proactive detection, interpretability, and user interaction, the proposed system is designed to improve both technical reliability and practical usability in a real-world elderly care environment [25]. As summarised in Table 1, existing approaches either focus solely on post-fall detection or lack personalization and human-centred interaction. In contrast, the proposed system uniquely integrates proactive fall risk prediction, online personalization, and human-in-the-loop confirmation within a real-time, non-wearable framework [26].

Table 1: Comparative analysis of existing fall detection and fall risk prediction systems

Study Category	**Sensing Modality**	**Detection Type**	**Personalization**	**Real-Time Operation**	**Human-in-the-Loop**	**False Alarm Mitigation**	**Deployment Cost**	**Privacy Concerns**	**Scalability**	**Limitations**

Thermal imaging-based methods	Depth sensor-based methods	Skeleton-based vision methods	Vision-based deep learning methods	Vision-based silhouette methods	Smartphone-based detection	Wearable sensor-based methods
Thermal camera	Depth camera (e.g., Kinect)	RGB / RGB-D camera	RGB / RGB-D camera	RGB camera	Built-in IMU (accelerometer, gyroscope)	Accelerometer / Gyroscope
Post-fall detection	Post-fall detection	Post-fall detection	Post-fall detection	Post-fall detection	Post-fall detection	Post-fall detection
No	No	No	Implicit (data-driven)	No	Limited	Limited (manual calibration)
Yes	Yes	Yes	Often limited	Yes	Yes	Yes
No	No	No	No	No	No	No
Temperature pattern analysis	Shape + distance thresholds	Fixed posture thresholds	Learned the decision boundary	Heuristic filtering	Threshold + activity filtering	Threshold tuning
High	Medium–High	Medium	High	Low	Low	Medium
Low	Low–Medium	Medium–High	High	High	Medium	Low
Moderate	Low	Moderate	High	Moderate	High	Moderate
Lower resolution; affected by ambient temperature	Limited range; sensitive to sensor placement	High false alarms during daily activities	Data-intensive; limited interpretability; poor generalization	Sensitive to lighting, occlusion, and background changes	Depends on phone carrying; inconsistent placement	Requires user compliance, discomfort, and battery dependency

Proposed method	Multi-sensor fusion methods	Pre-fall risk assessment (non-vision)	Ambient sensor-based methods	Radar-based monitoring systems
Monocular RGB camera	Combination (vision + wearable + ambient)	Motion capture / Depth sensors	Pressure sensors, floor sensors	RF/mm Wave radar
Pre-fall risk prediction	Pre- and post-fall detection	Pre-fall prediction	Post-fall detection	Pre- and post-fall detection
Online personalized baseline learning	Moderate	Limited	No	Limited
Yes	Often limited	Often offline	Yes	Yes
Yes (audio + gesture)	Rare	No	No	No
Temporal instability window + user confirmation	Sensor fusion algorithms	Controlled environments	Event-triggered sensing	Signal pattern filtering
Low	High	High	Medium	High
Medium (privacy-aware design)	Medium	Low	Low	Low
High	Low–Moderate	Low	Low	High
Designed primarily for indoor environments	Complex integration; high cost	Requires specialized hardware and a controlled setup	Limited coverage; installation complexity	Expensive hardware; complex signal processing

2.1. System Architecture

The proposed system is a real-time, non-wearable fall-risk monitoring framework for indoor environments. It operates using a single monocular RGB camera and follows a modular pipeline architecture to ensure computational efficiency, interpretability, and ease of deployment. An overview of the system architecture is illustrated in Figure 1 (to be included), and the main functional modules are described below.

2.2. Overview of the Processing Pipeline

The system architecture consists of six principal modules arranged in a sequential yet event-driven manner: (1) video acquisition, (2) skeletal pose estimation, (3) posture feature extraction, (4) personalized baseline learning and fall risk estimation, (5) human-in-the-loop interaction, and (6) alert escalation and logging. Video frames are processed continuously in real time, enabling uninterrupted fall risk monitoring during daily activities. Raw RGB video streams captured by the camera are first forwarded to the pose estimation module. Extracted skeletal data are then transformed into posture-related features, which are used to compute continuous fall risk scores. When elevated risk persists within a defined temporal window, the system activates the interaction module before escalating an emergency alert.

2.3. Video Acquisition and Pose Estimation

The video acquisition module captures RGB frames at a fixed frame rate suitable for real-time operation. Each frame is processed independently to avoid latency accumulation. Skeletal pose estimation is performed using MediaPipe Pose, which provides 2D coordinates and confidence scores for the major joints of the human body. The use of skeletal representations significantly reduces sensitivity to background clutter, lighting variations, and appearance changes, while preserving essential posture and motion information. Only upper-body and torso-related joints are retained for further analysis, as these are most relevant to balance and instability assessment. This method reduces the system's reliance on environmental factors, which can fluctuate significantly in actual homes (real-time applications). Even if the backdrop isn't clean or the lighting isn't ideal, the system model can still concentrate on body motion. This ensures steady performance without the requirement for controlled environments.

2.4. Posture Feature Extraction

The posture feature extraction segment computes posture descriptors that indicate body orientation and stability, relying on the anticipated skeletal joints. The torso's direction is determined by the positions of the shoulder and hip joints relative to one another, allowing us to assess the degree of tilt from vertical. Researchers monitor changes in joint positions and torso angles over time to observe dynamic instability. The advantages of posture feature extraction included reduced computational power, ease of grasp, and a direct connection to balancing the model. These features help to improve the decision-making capacity of real-time fall risk events without complex machine learning models. These features are also easy to interpret when analyzing system output. Instead of relying on complex hidden patterns, changes in posture can be understood directly, making debugging and improvement easier. This simplicity is helpful when the system needs to be tested or explained in real-world use.

2.5. Personalized Baseline Learning and Fall Risk Estimation

In real time, personalized baseline learning models account for differences in a person's standing and movements. It is closely monitoring how people move and stand in day-to-day life. The System Updates posture using recent data. It helps adjust quirks without requiring data to enhance performance. It always supports risk estimation by comparing the posture with the current posture to create an accurate learned baseline. Parallelly, persons' torsos were oriented, and if they were getting a risk score from the system model. Here, the model mainly aims to normalize the score; it doesn't say 'fall' or 'not', it keeps monitoring the risk level. The risk score indicates how stable a person was over time during these daily activities. The system checks a person's posture and movement. These help determine a person's stability. This gradual approach makes the system more personalized over time. It does not rely on pre-trained assumptions; instead, it adapts to actual usage. Because of this, it becomes more accurate for each individual and avoids general mistakes that can happen in fixed models.

2.6. Temporal Instability Observation Window

In real time, some activities lead to temporary elevations due to transient posture changes during voluntary movements, without indicating an actual unstable or hazardous situation. To address the problem, a temporal instability window is introduced, requiring the elevated risk to persist for a predefined duration before triggering the system model. Most of the false alarms were caused by short-term fluctuations that were suppressed, leaving only sustained instability patterns; these were considered potential pre-fall events. Maintaining the sensitivity and robustness of the observation window is treated as a temporal filter. This idea helps the system avoid overreacting to small or harmless movements. In daily life, people often make quick, low-risk adjustments. By waiting for consistency over time, the system becomes more stable and less sensitive to noise.

2.7. Human-in-the-Loop Interaction Module

Human verification is performed when the system detects a sustained high-risk situation before sending an alert. The monitored individual is informed of the detected risk via a voice prompt and waits for a response: raising the right hand indicates that the user is safe and doesn't need help. Also, this interaction with the older adults has a time limit; if the system detects the hand gesture within that time, it cancels the alert and returns to the monitoring stage. Having people involved in this system makes decisions, which ideally reduces false alarms. This interaction with the model incorporates a human aspect into the system, which is crucial for practical application. Rather than relying solely on automation, it enables the user to engage, giving the system a more supportive rather than dominant feel. This may enhance acceptance, particularly among older users.

2.8. Alert Escalation and Data Logging

An emergency notification is dispatched to caregivers or monitoring services as required if the system fails to receive confirmation from the user within the designated timeframe. The system monitors all activities throughout this situation. All specifics, such as skeletal information, fall risk scores, and events during interactions with timestamps, are stored in CSV files, facilitating validation and review of the system's performance, adjusting settings, and ensuring experiments can be replicated. Maintaining this information simplifies the analysis of the system's performance over time. It also helps identify trends or recurring problems, which can enhance performance.

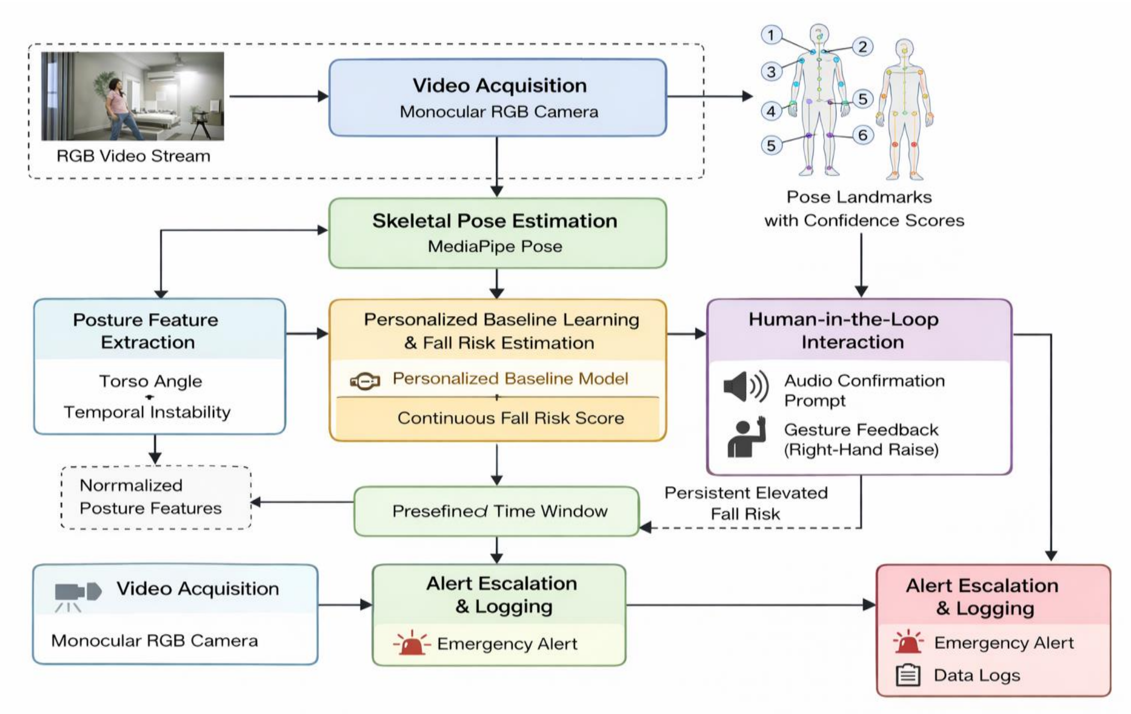

Figure 1: Overall architecture of the proposed vision-based fall risk prediction system

Ultimately, this data supports both system development and real-world monitoring. Figure 1 shows our vision-based system that predicts fall risk. This system is built to work step by step over time. It uses cameras to assess a person's posture and determine their fall risk. The goal of this vision-based system is to keep adults safe indoors. As seen in Figure 1, a regular video camera takes video in time. Then the MediaPipe Pose tool analyzes the video to find the body positions researchers need to study posture and movement. This information feeds into a part of the system that learns each person's normal posture. This helps the vision-based system adjust for the differences between people. Next, the vision-based system's fall-risk component analyses the person's body posture. It calculates a score indicating the likelihood that the person will fall.

This score is based on how much the person's torso is leaning and how unsteady they are on their feet. The vision-based system does not issue alerts when it detects rapid movement. Instead, it waits for a while to see if the person is unstable. If the score exceeds a threshold, the vision-based system asks the person if they are okay with using a voice prompt. The vision-based system aims to forecast fall risk and ensure individuals' safety. The individual may respond with a hand posture. If the system doesn’t receive a response promptly, it triggers an alert for assistance. The system also stores performance data in a CSV file for evaluation. To provide a clear and structured representation of the proposed fall risk prediction framework, the overall procedure is summarised in Algorithm 1. The algorithm outlines the key steps involved in real-time posture analysis, risk estimation, and alert generation.

Algorithm 1: Proactive Fall Risk Prediction System

Input: Continuous RGB video stream

Output: Real-time fall risk score Rt and alert notification

1: Initialize system parameters (Tr, Tw, μθ, μσ, α, w1, w2)
These parameters define the system's sensitivity and determine how quickly it adapts to user behavior.
2: while (video stream is active) do=
3: Capture current frame Ft
Frames are processed continuously to ensure smooth real-time monitoring.

4: Perform pose estimation and extract skeletal joints $\boldsymbol{J_t}$
Joint coordinates are obtained while ignoring low-confidence detections.
5: Select the shoulder and hip joints
These joints provide a stable representation of torso orientation.
6: Compute torso vector $\boldsymbol{V_t}$and torso angle $\boldsymbol{\theta_t}$
The angle indicates how much the body deviates from an upright position.
7: Update the sliding window with $\boldsymbol{\theta_t}$
Maintains recent posture values for short-term temporal analysis.
8: Compute temporal instability $\boldsymbol{\sigma_t}$
Captures fluctuations in posture that may indicate imbalance.
9: Update personalized baseline using exponential moving average
Gradually learns the normal posture pattern of the individual.
10: Compute normalized deviations ($\boldsymbol{d_{\theta,t}}$ and $\boldsymbol{d_{\sigma,t}}$)
Measures how far the current posture differs from the learned baseline.
11: Calculate fall risk score $\boldsymbol{R_t}$
Combines both deviation and instability into a continuous risk measure.
12: if ($\boldsymbol{R_t}$exceeds threshold $\boldsymbol{T_r}$for duration $\boldsymbol{T_w}$) then
Ensures that only sustained abnormal behavior is considered.
13: Trigger voice-based confirmation
The system prompts the user to confirm their safety.
14: if (gesture detected within time limit) then
15: Cancel alert and continue monitoring
Avoids unnecessary emergency notifications.
16: else
17: Send emergency alert to caregiver
Indicates a potential fall risk situation.
18: end if
19: end if
20: Log posture data, risk score, and timestamps
Enables further analysis and performance evaluation.
21: end while

Algorithm 1 provides a structured presentation of the suggested framework, encompassing all crucial phases from risk assessment and alarm creation to skeleton extraction. The system can operate effectively in real time while remaining resilient to brief posture changes, thanks to a combination of temporal analysis and customised baseline adaptation.

3. Methodology

Methodological details of the proposed real-time fall risk prediction system. The approach aims to ensure interpretability, real-time applicability, and user personalization, while sidestepping reliance on wearable devices or resource-intensive learning models. The main goal is to consistently assess fall risk by analyzing posture changes and to incorporate human validation before increasing alert severity. This method allows the system to operate silently in the background, requiring user attention. It doesn't require the individual to wear anything or constantly engage with devices, making it more appropriate for everyday use, particularly for older adults. The aim is to maintain simplicity and practicality to keep monitoring manageable. Another aspect is that individuals do not move in identical manners.

What appears normal to one person may seem strange to another. As a result, the system does not adhere to rigid established rules. Rather, it slowly learns what is typical for each person and utilizes that as a benchmark. This enhances the reliability of the results and decreases the likelihood of false alerts. The approach also minimizes intensive calculations whenever feasible. Because the system is anticipated to operate continuously, it must be efficient. By focusing solely on key posture-related aspects, it can deliver rapid results. This is crucial in scenarios where even minor delays can impact response time. The approach aims to harmonize precision, straightforwardness, and accessibility. It is designed for use in actual homes with minimal setup, while still delivering consistent, significant, and reliable outcomes over time.

3.1. Skeletal Representation and Joint Selection

Given an input RGB video stream, skeletal pose estimation is performed for each frame using MediaPipe Pose. The pose estimator outputs two-dimensional coordinates and confidence scores for a predefined set of body joints. Let:

$$\mathbf{J}_t = \{(x_{i,t}, y_{i,t})\}_{i=1}^{N} \quad (1)$$

Denote the set of detected joint coordinates at time t, where N is the number of detected joints. To make the system more stable and faster, only joints that affect posture stability are kept. The shoulder and hip joints are chosen to illustrate how the torso is oriented, as the torso's angle is closely related to balance and the risk of falling. To reduce noise caused by occlusion or detection errors, joints with confidence scores below a threshold are discarded. In this case, Equation (1) lists all the body joints found at that time. The horizontal (x) and vertical (y) positions of each joint in the image define it. In simple terms, it's like making a map of the human body that shows important points that can be followed over time. This mathematical representation helps turn pictures into structured data that can be easily processed. Instead of dealing with raw images, the system operates on coordinates, which speeds up and stabilizes computations.

3.2. Torso Angle Estimation

Torso orientation is modeled using the relative positions of the shoulder and hip joints. Let $(x_{s,t}, y_{s,t})$and $(x_{h,t}, y_{h,t})$denote the midpoints of the left–right shoulder pair and left–right hip pair at time t, respectively. The torso vector is defined as:

$$\mathbf{v}_t = (x_{s,t} - x_{h,t},\ y_{s,t} - y_{h,t}) \quad (2)$$

The torso angle $\theta_t$with respect to the vertical axis, it is computed as:

$$\theta_t = \arctan\left(\frac{x_{s,t} - x_{h,t}}{y_{h,t} - y_{s,t}}\right) \quad (3)$$

This angle provides an interpretable measure of body inclination. Large deviations from the upright posture are indicative of instability or loss of balance. Equation (2) defines a vector that connects the hip position to the shoulder position. This vector shows the direction the upper body is oriented. If the person is standing straight, this vector will be almost vertical. Equation (3) then converts this vector into an angle value. The arctangent function is used to calculate the angle of tilt of the torso from the vertical. A small angle means the person is upright, while a larger angle indicates leaning or bending. This transformation from position to angle makes it easier to interpret posture consistently across different frames and individuals.

3.3. Temporal Instability Measurement

Static posture deviation and temporal instability are analyzed to capture deterioration in dynamic balance. Temporal instability is defined as the short-term variability of the torso angle over a sliding window of length W. Let $\{\theta_{t-W+1}, \dots, \theta_t\}$denote the torso angles within the window. Temporal instability is quantified as:

$$\sigma_t = \sqrt{\frac{1}{W}\sum_{k=t-W+1}^{t} (\theta_k - \bar{\theta}_t)^2} \quad (4)$$

Where $\bar{\theta}_t$is the mean torso angle within the window. This measure captures oscillatory or erratic posture behavior that may precede a fall, even when absolute angle deviations were moderate. Equation (4) helps to find the standard deviation of the torso angle over a short period of time. To put it simply, it shows how much the angle is changing from one frame to the next. If σ_t is small, it means the person's posture is stable and changing little. A big value, on the other hand, indicates that movements occur often or suddenly, which may indicate that the person is off balance. This helps the system figure out not only where the body is, but also how steady or unstable the movement is over time.

3.4. Online Personalized Baseline Learning

Fall risk assessment can be significantly affected by intersubject variation in posture and movement patterns. The suggested approach uses an online, personalized baseline posture model to address this problem. Baseline statistics for torso angle and temporal instability were updated during daily activities. Let $\mu_\theta$and $\mu_\sigma$denote the baseline mean values. These parameters were kept updated by an exponential moving average:

$$\mu_t = \alpha\mu_{t-1} + (1-\alpha)x_t \quad (5)$$

Where $x_t$represents the current observation and $\alpha \in (0, 1)$controls the adaptation rate. Individual posture and characteristics have evolved through the online wearing technique. Equation (5) illustrates how the system updates the baseline using recent observations. It employs a weighted average, giving more weight to recent readings, rather than storing all historical data. The parameter α determines the system's rate of adaptation. A smaller value of α permits faster adaptation, whereas a greater number indicates slower changes. The system can learn new patterns while maintaining stability through this balancing. This technique guarantees that the baseline accurately depicts the user's posture throughout time, free from sudden shifts.

3.5. Continuous Fall Risk Scoring

The fall risk is represented as a scalar value. To always normalize the posture deviation and temporal instability, even at each step:

$$\boldsymbol{d_{\theta,t}} = \frac{|\theta_t - \mu_\theta|}{\mu_\theta}, \boldsymbol{d_{\sigma,t}} = \frac{|\sigma_t - \mu_\sigma|}{\mu_\sigma} \tag{6}$$

The overall fall risk score $\boldsymbol{R_t}$:

$$\boldsymbol{R_t = w_1 d_{\theta,t} + w_2 d_{\sigma,t}} \tag{7}$$

Where $\boldsymbol{w_1}$, $\boldsymbol{w_2}$: weighting coefficients. This formulation allows smooth risk evolution monitoring and provides early identification of pre-fall conditions. Equation (6) calculates the difference between the current posture and instability and the baseline. It normalizes the difference so that the values remain comparable. Equation (7) combines these normalized values into a single risk score. The weights w_1 and w_2 determine the importance given to posture deviation and instability. This combined score provides a simple yet effective way to continuously monitor fall risk, rather than making sudden decisions.

3.6. Temporal Instability Observation Window

To reduce the false alarms caused by transient or voluntary movements, a temporal instability observation window is applied to the risk score. A fall risk alert is considered only if $\boldsymbol{R_t}$exceeds a predefined threshold $\boldsymbol{T_r}$continuously for a duration $\boldsymbol{T_w}$. It doesn’t trigger unnecessary alerts based on persistence criteria, ensuring that live posture deviations are detected and improving operational robustness.

3.7. Human-in-the-Loop Confirmation Mechanism

The system feels more useful in daily use thanks to this confirmation step. People sometimes make abrupt or strange motions in real life that are not harmful, and these could be readily misinterpreted as fall hazards without verification. Giving the user an easy way to reply reduces the system's intrusiveness and helps prevent needless worry. For older persons who might not feel comfortable using digital gadgets, the gesture-based interface is straightforward to understand. Raising a hand is a natural gesture that makes it easier for the user to understand what the system expects when paired with an audible cue. This results in a more seamless and user-friendly overall experience. Over time, the system can also observe how often alerts are confirmed or cancelled. This can give useful feedback about how well the system is performing and whether adjustments are needed. In this way, the confirmation mechanism not only reduces false alarms but also helps the system improve gradually.

3.8. Alert Escalation Strategy

When someone failed to answer, this escalation stage becomes crucial. It can be a sign that the person truly needs assistance if there is no response within the allotted period. In these situations, the system doesn't wait any longer and sends an alarm, which may be crucial in preventing a serious outcome. This method permits adaptability based on the system's location. Alerts can be sent, for instance, to emergency contacts, family members, or carers. This ensures someone is promptly alerted and able to act when necessary, demonstrating their ability to balance caution and pragmatism. It can be annoying to receive too many notifications, but it is even more dangerous to miss an actual emergency. By waiting briefly for confirmation and acting only if needed, the system handles both situations reasonably.

3.9. Data Logging and Analysis Support

The retained data can also be beneficial beyond mere system functionality. Analyzing trends over time provides insight into how an individual’s mobility or stability has evolved. Such information can help identify gradual deterioration or improvement in physical condition. It also simplifies revisiting previous occurrences when necessary. For instance, when a fall or near-fall incident occurs, the recorded information can help clarify what happened before it. This can enhance systems, assess medical conditions, and maintain well-structured data, which aid the continued advancement of the system. It serves to evaluate new concepts, enhance current techniques, or assess performance against alternative methods. This enhances the system's adaptability and long-term utility.

4. Results

Experimental results obtained with the help of the real-time indoor monitoring by evaluating the poster in dynamics, risk evaluation, and false-alarm suppression.

4.1. Posture Dynamics and Instability Analysis

Figure 2 and Table 2 show how the estimated torsion angle changes over time frame by frame. In most cases observed, the torso angle is centred on the person's personalised baseline, reflecting how they perform their daily tasks. Some short-term changes matched movements like bending and reclining.

Table 2: Torso angle variation data

Timestamp	Torso Angle (°)
1771342351	172.30
1771342351	173.01
1771342351	173.44
1771342351	173.82
1771342352	175.55
1771342352	176.12
1771342352	178.85
1771342352	177.17
1771342352	176.52

In contrast, sustained deviations observed in later segments indicate periods of postural instability that may precede falls. It should be evident that the extent and duration of these deviations are essential in differentiating normal activities from potential risk conditions. Short, controlled deviations usually return to baseline quickly, but unstable movements tend to last longer, indicating the person is losing their balance.

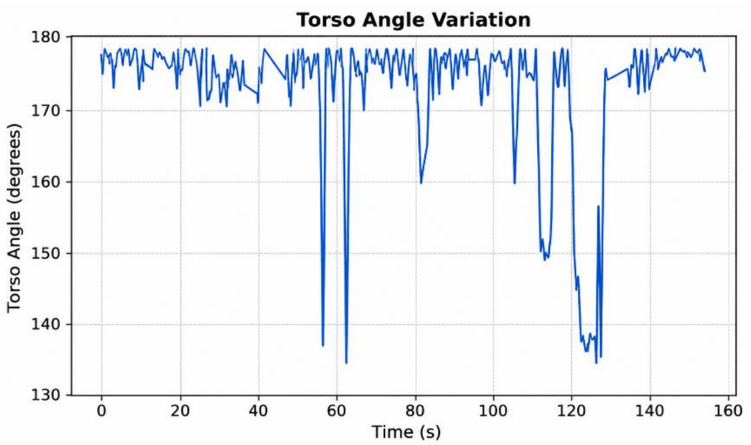

Figure 2: Torso angle variation

The baseline alignment's consistency across most time intervals indicates that the personalized modeling approach accurately captures individual posture patterns and reliably identifies abnormal deviations. The postural instability metric, shown in Figure 3 and Table 3, shows changes in torso angle from frame to frame. The instability signal remains close to zero and shows little variation when the posture is stable.

Table 3: Temporal instability measurements

Time	Instability (σt)
1771342351	-0.99
1771342351	1.13
1771342351	0.48
1771342351	0.54

1771342352	3.38
1771342352	4.03
1771342352	6.60
1771342352	4.15
1771342352	2.70

When the posture is unstable, the signal shows large spikes and greater variation. The instability pattern matches the long-term torsional deviation, indicating a dynamic imbalance.

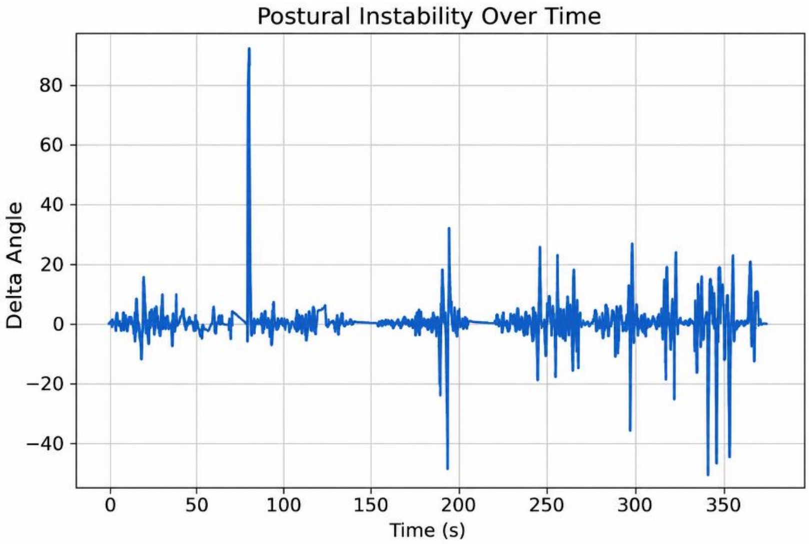

Figure 3: Corresponding postural instability

The presence of repeated spikes in the instability signal indicates that an imbalance persists during unstable periods, not just during sudden changes. The clear distinction between stable and unstable areas further enhances the system's ability to accurately assess fall risk.

4.2. Fall Risk Score Evolution

Figure 4 shows that the system has been monitoring the continuous fall risk score. During stable posture and brief voluntary movements, the risk remains the same.

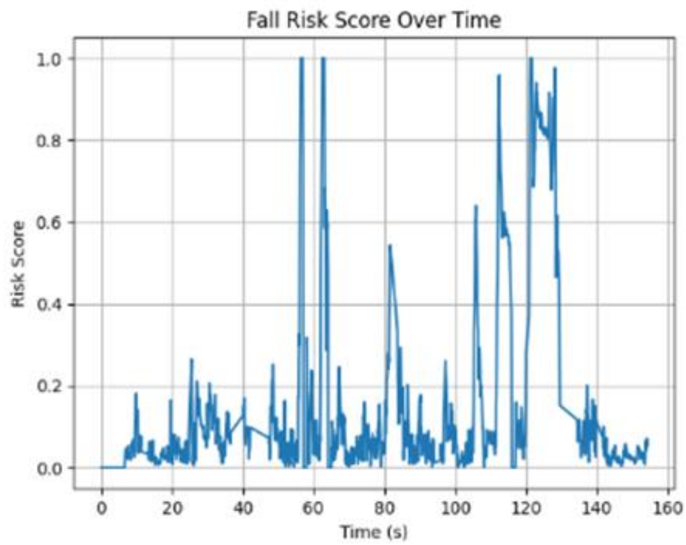

Figure 4: Fall risk score

Prolonged torso inclination and increased instability during sustained imbalance lead to a progressive rise in fall risk approaching peak levels. The temporal alignment between Figures 2 to 4 confirms that continuous instability, not single posture alterations, has become the cause of higher risk. The transition from stable posture to high-risk conditions is demonstrated in Table 4. At first, there is little instability and little risk, and the torso angle stays around the upright position.

Table 4: Fall risk score and system response

Time	Torso Angle (°)	Instability (σt)	Risk Score (Rt)	Status
1771342378	177.81	-1.28	0.025	Safe
1771342379	173.61	-5.06	0.160	Safe
1771342379	172.98	-5.28	0.175	Safe
1771342379	172.98	-4.85	0.169	Safe
1771342379	170.76	-7.08	0.243	Safe
1771342379	167.97	-9.00	0.325	Safe
1771342379	168.60	-7.96	0.298	Safe
1771342379	161.11	-15.87	0.554	Unstable
1771342379	156.61	-21.69	0.721	High Risk
1771342379	152.59	-26.57	0.867	High Risk
1771342379	150.08	-28.66	0.945	High Risk

The risk score gradually increases as instability rises sharply due to a significant deviation in posture. The efficacy of the suggested fall risk prediction model is ultimately confirmed when the system detects high-risk situations in which both torso deviation and instability reach critical levels.

4.3. Relationship Between Torso Angle and Fall Risk

Figure 5 shows the relationship between torso angle and fall risk score. The torso angle value is close to the baseline, resulting in a low-risk score, while increasing the torso value leads to progressively higher risk scores.

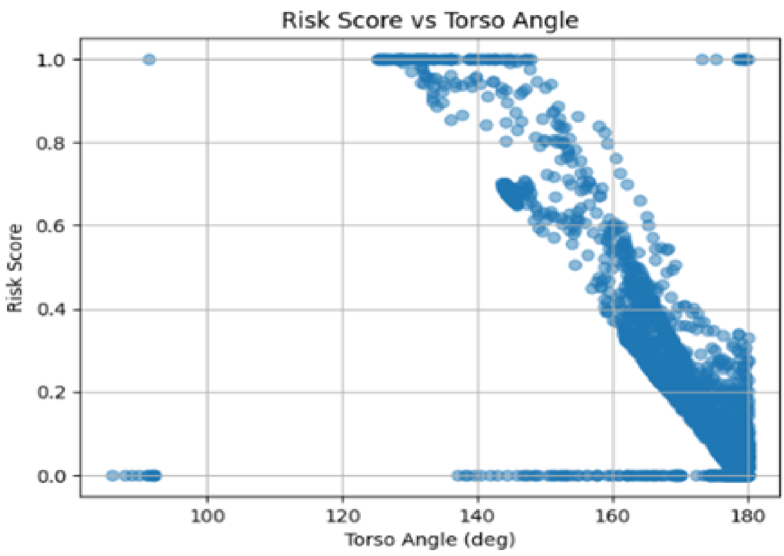

Figure 5: Torso angle and full risk

The data scatter indicates that large posture deviations alone do not necessarily trigger high risk unless accompanied by sustained instability.

4.4. False Alarm Reduction Performance

The effectiveness of the human-in-the-loop confirmation mechanism, the False Alarm Reduction Rate (FARR), is defined as the proportion of high-risk events that are canceled due to user feedback before alert escalation. The FARR is computed as:

$$\mathbf{FARR} = \frac{\boldsymbol{N}_{\text{cancelled}}}{\boldsymbol{N}_{\text{high-risk}}} \quad (8)$$

Where $\boldsymbol{N}_{\text{cancelled}}$ denotes the number of high-risk events canceled by the person, and $\boldsymbol{N}_{\text{high-risk}}$ represents the total number of detected high-risk events that occurred. The system achieved a False Alarm Reduction Rate 1.9% of potential emergency alerts were successfully suppressed. These canceled events corresponded to voluntary or recoverable posture deviations that did not require intervention. It clearly shows that the human-in-the-loop mechanism effectively reduces unnecessary alerts.

4.5. Effectiveness of Temporal Observation Window

To enhance the robustness, the temporal instability observation window has placed the majority of the poster deviation and instability spikes visible in Figures 2 and 3. It doesn't result in the sudden change or rise of the risk peak in Figure 4. It indicates the effective model. By requiring sustained elevated risk before alert activation, the system achieves a balance between sensitivity and false-alarm suppression, making it suitable for real-world indoor monitoring.

5. Discussion

Table 5 compares the proposed system with other fall-monitoring methods, focusing on their configuration. Methods that use sensors are effective in some situations, but they require people to use them correctly, which can be a problem, especially with older people. On the other hand, methods that use vision and deep learning do not require wearable sensors, but they can be slow and use a lot of computer power, making them hard to run in real time on less powerful systems. Also, these methods can be hard to understand. Do not give users too much information, as this can make them distrust you. Most vision-based fall detection systems typically rely on predefined rules and thresholds. These systems are simple to set up. They are not very flexible and can give false alarms when people do normal things, like sitting down quickly or bending. This shows that researchers need a system that can adapt to the user. The proposed system solves these problems by using a camera and analyzing the person's skeleton in real time. It differs from methods in that it can learn a person's normal posture and movement patterns over time. This helps the system distinguish between unsafe behaviors. The system also uses a scoring system to detect falls based on a person's movement, helping it identify subtle signs of instability. To make the system more reliable, it asks the user to confirm if they are okay before sending an alert. This helps reduce alarms. By doing this, the system can minimize notifications. The proposed system is very effective at reducing alarms, achieving a rate of about 1.9%, indicating it works well in real-world situations. The proposed system is superior to fall-monitoring methods because it uses a camera to analyze a person's movements in real time. The proposed system is effective at detecting falls and reducing false alarms, making it a good choice for people who need a reliable fall monitoring system.

Table 5: Execution on-level comparison

Aspect	Wearable-Based Methods	Vision-Based DL Methods	Traditional Vision-Based Methods	Proposed System
Sensor modality	Wearable inertial sensors	RGB / RGB-D cameras	RGB cameras	Monocular RGB camera
User compliance required	High	None	None	None
Computational complexity	Low–moderate	High	Low	Low
Real-time feasibility	Yes	Hardware-dependent	Yes	Yes
Detection focus	Post-fall detection	Mostly post-fall	Post-fall	Pre-fall risk prediction
Personalization strategy	Manual calibration	Dataset-driven	None	Online baseline learning
Decision mechanism	Threshold-based	Model inference	Threshold-based	Continuous risk scoring
Temporal modeling	Limited	Implicit	Limited	Explicit temporal instability window
False alarm handling	Limited	Dataset-dependent	Poor	Temporal filtering + user confirmation
False alarm reduction metric	Rarely reported	Inconsistent	Not addressed	FARR = 0.019 (1.9%)

Human-in-the-loop interaction	Not supported	Not supported	Not supported	Audio + gesture-based feedback
Model interpretability	High	Low	High	High
Deployment suitability	Moderate	Limited	Moderate	High

6. Conclusion and Future Work

This paper discusses the development of a fall-risk monitoring application to enhance the safety of individuals, particularly older adults, in indoor environments. Using a camera-based method, the system continuously monitors and analyses the person's movements, paying special attention to their balance and posture, and even uses a personalized baseline for each person. The system doesn't just look for falls after they happen; it also tries to predict them by watching for these visual clues. One of the best things about the system is that it can learn a unique model for each person. The system learns and distinguishes between normal and abnormal behaviour for everyone, rather than relying on general thresholds. These patterns may indicate a greater risk of falling, such as instability, unusual postural deviations, or unpredictable movement dynamics. Because older people or people in indoor environments can move in many ways, this personalized learning method is more accurate and flexible, adapting to how the person reacts. To make the system even more reliable and reduce false alarms, it has a human-in-the-loop confirmation mechanism.

To keep the system model more accurate, the support system doesn't issue an alert immediately when it detects a potential fall risk. Instead, it plays an audio prompt and waits for the user's response. It sends an emergency alert if there is no response after a specified period. This reduces unnecessary alerts and makes users more confident in the system by letting them check whether they are safe or need help. Tests show that this system can detect fall risks in real time and reduces false alarms by a lot more than older methods. It doesn't just react to single moments; it tracks changes in posture over time, which helps it work more consistently. Such a safety monitoring system is very important for helping seniors live independently and get help when they need it. User interaction makes the system more useful and user-friendly by keeping the user involved in the monitoring process. Future improvements will make the system more robust in difficult situations, such as when many people are in the same area, lighting changes, or someone is partially hidden. Also, keeping data safe and private will still be a top priority. The system should also be tested in real homes over time to better understand how it performs in the real world. These features show that they will work well, be easy to use, and be flexible in the future.

Acknowledgement: The authors sincerely thank Rajalakshmi Engineering College, Rajalakshmi Institute of Technology, CMR Institute of Technology, Sathyabama Institute of Science and Technology, and the University of Edinburgh for their valuable support, resources, and academic environment that contributed to this research.

Data Availability Statement: The corresponding author can provide the data supporting the findings upon request.

Funding Statement: This research did not receive any specific funding.

Conflicts of Interest Statement: The authors declare that they have no financial, personal, or professional conflicts of interest that may have influenced the research, its outcomes, or the content of the manuscript.

Ethics and Consent Statement: This study was conducted in accordance with ethical research standards, and no human participants, personal data, or animal subjects were directly involved, formal ethical approval and consent were not required.

References

1. World Health Organization, “Falls,” *World Health Organization*, 2021. [Accessed by 12/03/2025].
2. C. Gangadhar, P. P. Roy, R. D. Kumar, J. V. N. Ramesh, S. Ravikanth, and N. Akhila, “Wearable sensor-based fall detection for elderly care using ensemble machine learning techniques,” *Measurement: Sensors*, vol. 39, no. 6, p. 101870, 2025.
3. M. Mubashir, L. Shao, and L. Seed, “A survey on fall detection: Principles and approaches,” *Neurocomputing*, vol. 100, no. 1, pp. 144–152, 2013.
4. E. Alam, A. Sufian, P. Dutta, and M. Leo, “Vision-based human fall detection systems using deep learning: A review,” *Computers in Biology and Medicine*, vol. 146, no. 7, p. 105626, 2022.
5. A. Benkaci, L. Sliman, and H. N. Dellys, “Vision-based Human Fall Detection Systems: A Review,” *Procedia Computer Science*, vol. 241, no. 8, pp. 203–211, 2024.

6. N. Noury, A. Fleury, P. Rumeau, A. K. Bourke, G. O. Laighin, V. Rialle, and J. E. Lundy, "Fall detection-principles and methods," in *2007 29th annual international conference of the IEEE engineering in medicine and biology society*, Lyon, France, 2007.
7. F. Bagala, C. Becker, A. Cappello, L. Chiari, K. Aminian, J. M. Hausdorff, W. Zijlstra, and J. Klenk, "Evaluation of accelerometer-based fall detection algorithms on real-world falls," *PloS One*, vol. 7, no. 5, pp. 1–9, 2012.
8. S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *Journal of neuroengineering and rehabilitation*, vol. 9, no. 1, p. 21, 2012.
9. T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer vision and image understanding*, vol. 81, no. 3, pp. 231–268, 2001.
10. Z. Cao, G. Hidalgo, T. Simon, S. E. Wei, and Y. Sheikh, "Openpose: Realtime multi-person 2d pose estimation using part affinity fields," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 1, pp. 172–186, 2021.
11. A. Núñez-Marcos, G. Azkune, and I. Arganda-Carreras, "Vision-based fall detection using convolutional neural networks," *Wireless Communications and Mobile Computing*, vol. 2017, no. 1, pp. 1–16, 2017.
12. J. M. Hausdorff, "Gait dynamics, fractals and falls: Finding meaning in the stride-to-stride fluctuations of human walking," *Human Movement Science*, vol. 26, no. 4, pp. 555–589, 2007.
13. Y. Nizam, M. N. H. Mohd, and M. M. A. Jamil, "Development of a User-Adaptable Human Fall Detection based on Fall Risk Levels using Depth Sensor," *Sensors*, vol. 18, no. 7, p. 2260, 2018.
14. G. Mastorakis and D. Makris, "Fall detection system using Kinect's infrared sensor," *Journal of Real-Time Image Processing*, vol. 9, no. 3, pp. 635–646, 2014.
15. R. Igual, C. Medrano, and I. Plaza, "Challenges, issues and trends in fall detection systems," *Biomedical engineering online*, vol. 12, no. 1, p. 66, 2013.
16. A. L. Weng, M. Nysæther, M. Haddara, and M. Langseth, "User Satisfaction with Fall Detection Systems in Smart Elderly Care," *Procedia Computer Science*, vol. 256, no. 3, pp. 583–593, 2025.
17. J. Gutiérrez, V. Rodríguez, and S. Martin, "Comprehensive review of vision-based fall detection systems," *Sensors*, vol. 21, no. 3, p. 947, 2021.
18. X. Wang, J. Ellul, and G. Azzopardi, "Elderly fall detection systems: A literature survey," *Frontiers in Robotics and AI*, vol. 7, no. 6, p. 71, 2020.
19. Y. Wang and T. Deng, "Enhancing elderly care: Efficient and reliable real-time fall detection algorithm," *DIGITAL HEALTH*, vol. 10, no. 2, pp. 1–11, 2024.
20. W. S. Almukadi, F. Alrowais, M. K. Saeed, A. E. Yahya, A. Mahmud, and R. Marzouk, "Deep feature fusion with computer vision driven fall detection approach for enhanced assisted living safety," *Scientific Reports*, vol. 14, no. 9, pp. 1–17, 2024.
21. F. X. Gaya-Morey, C. Manresa-Yee, and J. M. Buades-Rubio, "Deep learning for computer vision based activity recognition and fall detection of the elderly: A systematic review," *Applied Intelligence*, vol. 54, no. 7, pp. 8982–9007, 2024.
22. T. Alanazi, K. Babutain, and G. Muhammad, "A Robust and Automated Vision-Based Human Fall Detection System using 3D Multi-Stream CNNs with an Image Fusion Technique," *Applied Sciences*, vol. 13, no. 12, p. 6916, 2023.
23. S. Li, L. Ding, and Y. Shi, "Research on fall detection based on vision and wearable devices," *Academic Journal of Computing & Information Science*, vol. 7, no. 11, pp. 1–10, 2024.
24. C. Kiranmai, B. Srivalli, C. H. Komali, G. Apurva, and B. S. Yesshaswi, "Visual Fall Detection Analysis through Computer Vision and Deep Learning–Technology Proposition," *International Journal of Recent Technology and Engineering*, vol. 13, no. 1, pp. 1–4, 2024.
25. M. A. Syamlan, A. Arifin, J. Pramudijanto, F. Arrofiqi, R. A. M. Noor, and A. S. F. Suhartono, "A Low-Cost Wearable System to Detect Fall and Non-Fall Activities for Elderly Individuals," *Journal of Novel Engineering Science and Technology*, vol. 5, no. 1, pp. 70–79, 2026.
26. E. R. Sykes, "Next-generation fall detection: Harnessing human pose estimation and transformer technology," *Health Systems*, vol. 14, no. 2, pp. 85–103, 2025.