TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 24, No. 6, December 2026, pp. 2145~2156
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v24i6.28141

Event-driven, privacy-preserving, edge-based multimodal fall detection for elderly care

Sufyan Yazid Abubakar[1], Rozeha Binti A. Rashid[1], Zakarie Abdi Mohamud[2]
[1]School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Johor, Malaysia
[2]Department of Telecommunication Engineering, Faculty of Engineering, Hormuud University, Mogadishu, Somalia

Article Info

Article history:

Received Jul 1, 2026
Revised Oct 5, 2026
Accepted Oct 11, 2026

Keywords:

Ambient assisted living
Edge computing
Fall detection
Internet of Things
Privacy preservation

ABSTRACT

Falls are a leading cause of injury-related hospitalization and mortality among older adults, and prompt detection is essential to prevent severe complications from delayed assistance. Existing approaches including wearable, vision-based, ambient, and multimodal systems address only part of this challenge: wearables require sustained user compliance, vision-based systems raise privacy and energy concerns through continuous camera operation, and multimodal systems often inherit these limitations through always-on visual pipelines. This paper proposes an event-driven, privacy-preserving multimodal fall detection system that integrates posture analysis, passive infrared (PIR) motion sensing, and audio distress detection on a resource-constrained edge device. Unlike conventional multimodal systems, the camera activates only when low-power sensors detect a potential event, while a training-free rule-based fusion strategy verifies falls and generates alerts, with all processing performed locally. Evaluated across 50 controlled trials, the system achieved 98.00% accuracy and 100% recall, delivering alerts within 9 s while reducing hourly energy consumption by 46.3% compared with continuous monitoring. These results demonstrate that the proposed system provides a reliable, privacy-preserving, and energy-efficient solution for edge-based smart-home elderly care.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Rozeha Binti A. Rashid
School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia
81310 Skudai, Johor, Malaysia
Email: rozeha@utm.my

1. INTRODUCTION

Falls are among the most serious health hazards affecting older adults and remain a leading cause of injury-related hospitalization, disability, and mortality. Most falls occur in domestic environments, where older adults spend most of their time [1], [2]. When a person is unable to rise without assistance after a fall, a situation commonly referred to as a long lie, the resulting immobility can lead to hypothermia, bronchopneumonia, dehydration, pressure ulcers, and post-fall anxiety [2], [3]. With more than 37.3 million falls requiring medical attention each year and the global elderly population continuing to grow, the demand for reliable, unobtrusive, and practical fall detection systems has become increasingly urgent [1], [4].

Recent advances in low-cost Internet of Things (IoT) hardware, embedded computing, and ambient sensing have enabled intelligent health monitoring without continuous human supervision [5], [6]. Ambient assisted living (AAL) systems provide context-aware and unobtrusive support for older adults, with fall detection remaining one of their most extensively studied applications [5]. Nevertheless, designing a fall detection system that simultaneously achieves high reliability, energy efficiency, privacy preservation, and practical deployment in residential environments continues to be a significant challenge [2], [7].

Existing fall detection approaches can be broadly classified into four categories: wearable, vision-based, ambient sensing, and multimodal systems. Each offers different trade-offs in terms of privacy, computational complexity, energy consumption, and robustness against false alarms. Wearable devices expose minimal privacy-sensitive information but rely on consistent user compliance, correct placement, and regular battery charging [8], [9]. Vision-based systems provide rich contextual information and generally achieve high detection accuracy; however, continuous camera operation raises privacy concerns and increases computational complexity, particularly when deep-learning models are employed [2], [7], [10]. Ambient sensing approaches, including passive infrared (PIR) and acoustic sensors, are non-intrusive, privacy-preserving, and energy-efficient, but they provide limited contextual information when used independently. For example, PIR-only systems often struggle to distinguish genuine falls from routine human movement [11], while a 12-week in-home evaluation of a commercial audio-based system reported a fall detection specificity of only 6.8%, with most false alarms triggered by ordinary household sounds rather than actual falls [12]. Multimodal systems overcome some of these limitations by combining complementary sensing modalities, typically achieving higher accuracy than any single modality alone. However, this improvement in robustness is usually gained at the cost of privacy and energy efficiency: most existing frameworks, including those described as edge-based, still rely on continuously operating cameras, thereby inheriting the same privacy exposure and energy costs as pure vision-based systems, while offering little advantage over ambient-only approaches in terms of power consumption [4], [13].

These limitations reveal a clear research gap. Although multimodal sensing improves robust detection, most existing systems continue to depend on continuous visual monitoring, which remains the primary source of privacy concerns and unnecessary energy consumption. Moreover, relatively few studies evaluate energy efficiency, duty-cycle behavior, or real-time latency under realistic smart-home conditions, limiting evidence of their practical deployment [14]. Event-driven sensing, in which low-power ambient sensors activate camera-based verification only when a potentially abnormal event is detected, remains relatively underexplored in multimodal fall detection despite its potential to reduce energy consumption while maintaining the robustness of multimodal sensing [15], [16].

To address these limitations, this paper proposes an event-driven, privacy-preserving multimodal fall detection system for smart-home elderly monitoring. The proposed framework integrates vision-based posture analysis, PIR motion sensing, and audio distress detection on a resource-constrained edge device. Low-power PIR and audio sensors operate continuously, activating the camera only when a potentially abnormal event is detected, thereby minimizing unnecessary visual acquisition while preserving the robustness of multimodal sensing. All data processing is performed locally, ensuring that no raw visual or audio data is transmitted beyond the edge device. The proposed framework demonstrates that event-driven activation and multimodal fusion are complementary rather than competing design strategies, illustrating how event-driven sensing, lightweight edge intelligence, and multimodal fusion can be effectively integrated into a privacy-preserving smart-home monitoring system. In doing so, the proposed approach directly addresses key challenges in IoT, edge computing, and AAL for elderly care. The main contributions of this work are summarized as follows: (i) event-driven activation: low-power PIR and audio sensors continuously monitored the environment and triggered the camera only when abnormal activity was detected, reducing energy consumption compared to always-on approaches; (ii) edge-based privacy preservation: all sensing and inference were executed locally on a Raspberry Pi 4, eliminating cloud transmission of raw sensor data and enabling low-latency detection; (iii) lightweight rule-based fusion: a computationally efficient rule-based mechanism combined visual posture, PIR motion, and audio distress evidence without requiring model training or labeled datasets; (iv) system-level evaluation: the system was evaluated across detection accuracy, response latency, and energy efficiency including a duty-cycle-based energy comparison between monitoring modes; and (v) practical alerting pipeline: an end-to-end alert mechanism delivered fall notifications through email, Telegram, and Favoriot IoT to support caregiver response in elderly care environments.

The remainder of this paper is organized as follows: section 2 describes the system design and methodology. Section 3 presents experimental results and discussion. Section 4 concludes the paper and outlines directions for future work.

2. METHOD

2.1. System architecture

The proposed system adopts a three-layer architecture comprising a sensing layer, an edge processing layer, and an alerting layer. These layers implement an event-driven monitoring pipeline that transitions from low-power sensing to event verification and alert generation while ensuring that raw sensor data are processed locally on the edge device. Figure 1 illustrates the overall system architecture and the corresponding data flow between its functional layers. The current implementation targets single-room

residential monitoring; extending the architecture to larger homes or multi-room, multi-resident environments would require additional sensing nodes and a coordination mechanism to distinguish concurrent events across locations. This limitation is discussed further as a direction for future work in section 3.7.

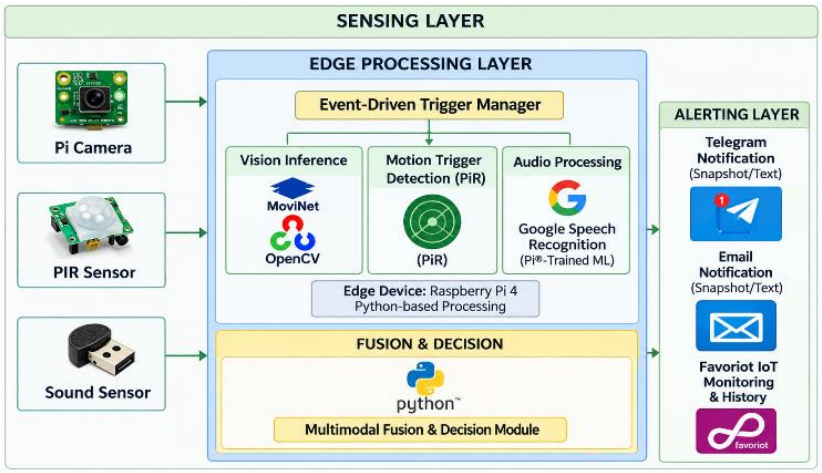


Figure 1. Overall architecture of the developed event-driven multimodal fall detection system, showing the sensing, edge processing, and alerting layers and the flow of data between them

The sensing layer integrates three complementary sensing modalities: a Pi Camera (IMX219) for vision-based posture analysis, a PIR motion sensor (HC-SR501) for detecting human movement and prolonged inactivity, and a universal serial bus (USB) microphone for capturing distress speech and impact sounds. By combining motion, visual, and acoustic information, the system improves fall detection reliability while reducing false alarms associated with individual sensing modalities. To ensure consistent experimental conditions, the camera and PIR sensor were installed at fixed positions and verified before data collection, as described in sections 2.3 and 2.4.

In the edge processing layer, all sensing data are processed locally on a Raspberry Pi 4, selected for its capability to perform real-time inference in resource-constrained environments. During Idle operation, an event-driven trigger manager continuously monitors the PIR sensor and microphone while keeping the camera inactive to reduce energy consumption. When an abnormal event is detected by either low-power sensor, the trigger manager activates the camera and initiates a short verification window, during which vision-based posture analysis and audio processing are performed concurrently. The outputs from the three sensing modules are then combined by a rule-based multimodal fusion module to determine the event priority and whether an alert should be issued. Once a fall is confirmed, the alerting layer notifies caregivers through Email and Telegram while recording structured event information on the Favoriot IoT platform for remote monitoring and historical event tracking. To preserve user privacy, only processed event information is transmitted from the edge device, whereas raw visual and audio data remain local. Table 1 summarizes the hardware and software components used in the proposed system.

Table 1. Hardware and software components of the developed fall detection system

Component	Model/tool	Function
Edge device	Raspberry Pi 4	Local edge processing and system control
Camera module	Pi camera (IMX219)	Vision-based posture analysis
Motion sensor	PIR sensor (HC-SR501)	Motion detection and event triggering
Audio sensor	USB microphone / sound sensor	Distress speech and impact detection
Programming language	Python 3.9	System implementation
Computer vision library	OpenCV	Image processing and frame handling
Pose estimation model	TensorFlow Lite MoveNet	Human pose estimation
Audio processing	Google speech recognition (speech recognition library)	Distress speech detection
Audio input library	Sounddevice	USB microphone audio streaming
PIR control library	General-purpose input/output (GPIO) zero (motion sensor)	GPIO-based PIR sensor interface
Camera library	Picamera2	Pi Camera frame capture and control
Numerical processing	NumPy	Keypoint array computation
Alert delivery	smtplib, requests	Email and Telegram notification dispatch
Notification platforms	Telegram, Email, Favoriot	Alert delivery and IoT monitoring

2.2. Event-driven operational workflow

The proposed system operates through five sequential states that together form the event-driven monitoring cycle: idle, triggered, verification, alert, and cooldown. Figure 2 illustrates the complete event-driven operational workflow and the transition conditions between these states, from continuous low-power monitoring in the idle state to sensor-triggered verification, priority-based alert generation, and the cooldown state before returning to idle.

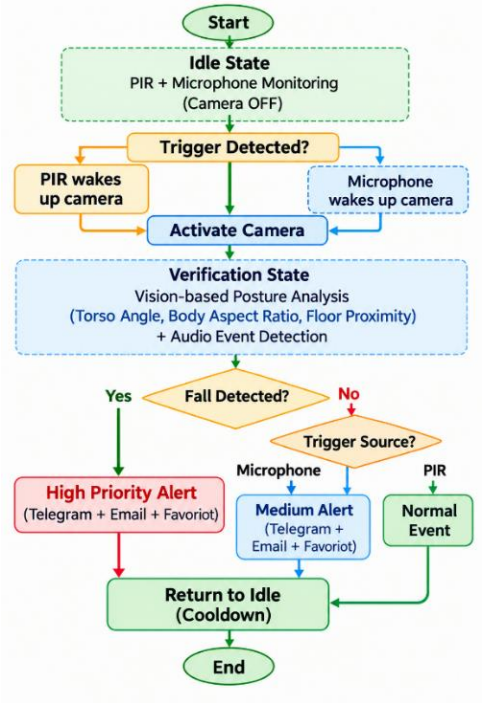


Figure 2. Event-driven operational workflow showing sensor-triggered verification, priority-based alert generation, and transitions between system states

The state transitions and their governing thresholds and timing parameters are summarized in Table 2, while the audio detection parameters are given in Table 3. Once a fall is confirmed, alerts are dispatched as described in section 2.7.

Table 2. Event-driven timing parameters

Parameter	Value	Description
Inactivity duration	15 s	Detects abnormal inactivity
Camera verification window	5 s	Duration for posture analysis
Impact/thud threshold	≈ −45 dBFS	Detects sudden impact sound
Cooldown period	10 s	Prevents repeated alerts

Table 3. Audio event detection parameters

Parameter	Value/method	Purpose
Distress keywords	Help, help me, oh my God, oh, ah	Identify verbal distress signals
Impact sound	≈ −45 dBFS	Detect sudden loud impact sounds
Speech processing	Google speech recognition + keyword matching	Convert speech to text and classify

2.3. Vision-based posture analysis

Vision-based analysis serves as the primary mechanism for confirming fall events. Instead of processing raw images directly, the proposed system employs the TensorFlow Lite MoveNet Lightning model to estimate human pose using 17 skeletal keypoints. This skeleton-based representation enables

posture analysis without storing or transmitting identifiable visual information, thereby preserving user privacy while supporting real-time inference on Raspberry Pi 4 [14].

To characterize fall-related postures, three complementary geometric features are extracted from the estimated skeleton: torso angle, body aspect ratio (AR), and floor proximity. Together, these features describe body orientation, body shape, and the relative position of the upper body with respect to the floor.

2.3.1. Torso angle (θ)

The torso vector is defined by the line connecting the mid-hip point (H_{mid}) and the mid-shoulder point (S_{mid}), where each midpoint is obtained by averaging the corresponding left and right body keypoints. The torso angle (θ) is then calculated relative to the vertical reference axis $v = (0,1)$:

$$\theta = \cos^{-1}\left(\frac{(S_{mid}-H_{mid})\cdot v}{\|S_{mid}-H_{mid}\|\|v\|}\right) \quad (1)$$

a posture is considered fall-like when the torso angle satisfies, $\theta \geq \theta_{th}$, where $\theta_{th} = 75^{\circ}$.

2.3.2. Body AR

The body AR is computed from the bounding box enclosing all detected skeletal keypoints:

$$AR = \frac{W}{H} \quad (2)$$

where, W and H denote the width and height of the bounding box, respectively. A posture is regarded as a fall candidate when $AR \geq AR_{th}$, where $AR_{th} = 1.2$. This threshold was empirically selected for the proposed camera configuration and is consistent with values reported for comparable threshold-based fall detection systems [17].

2.3.3. Floor proximity (d)

The normalized floor-proximity measure is computed as:

$$d = \frac{y_{floor}-y_{joint}}{H_{frame}} \quad (3)$$

where y_{floor} denotes the floor reference position, y_{joint} is the vertical coordinate of the evaluated upper-body joint, and H_{frame} represents the image height. The NearFloor condition is satisfied when the detected head or shoulder keypoints lie within a threshold of $d_{th} = 0.30$.

2.3.4. Fall detection logic

The extracted posture features are combined using the following rule-based decision:

$$Fall = [(\theta \geq \theta_{th}) \wedge NearFloor] \vee [(AR \geq AR_{th}) \wedge NearFloor] \quad (4)$$

The two criteria complement one another: the torso-angle criterion is generally more effective for detecting pronounced upper-body inclination, whereas the aspect-ratio criterion better captures horizontally elongated body configurations. Requiring the NearFloor condition in both cases suppresses false detections caused by activities such as sitting, bending, or kneeling.

2.3.5. Threshold calibration

The threshold values used in this study were empirically calibrated through preliminary experiments to achieve a balance between fall detection sensitivity and false-alarm suppression. Because the experimental dataset was limited, these values should be regarded as empirically calibrated rather than statistically optimized. A larger-scale statistical optimization, such as receiver operating characteristic (ROC)-based threshold analysis, is left for future work.

2.4. PIR motion and inactivity detection

The PIR sensor (HC-SR501) serves as the primary low-power trigger for the proposed event-driven monitoring system. Operating continuously throughout system operation, it detects changes in infrared radiation caused by human movement while consuming substantially less energy than continuous vision processing. The PIR sensor's underlying analog output can also support more advanced motion-classification techniques beyond simple binary triggering [18], though this was not explored in the present implementation.

To ensure reliable performance, the sensor's detection range and field of view were verified before experimentation under the fixed installation configuration used throughout this study.

In addition to detecting motion, the PIR sensor monitors prolonged inactivity following a motion event. If no further movement is detected within 15 s, the system interprets the inactivity as a potential abnormal event and activates the verification stage. This dual-function strategy enables the PIR sensor to provide an energy-efficient trigger for both movement and inactivity events and is consistent with ambient sensing approaches commonly adopted in smart-home elderly monitoring systems [19].

2.5. Audio distress and impact detection

The audio module provides a complementary sensing modality by detecting distress speech and impact sounds that may accompany a fall. A USB microphone continuously monitors the environment, processing audio in short overlapping windows to enable low-latency event detection. Captured speech is converted to text using Google speech recognition and matched against a predefined set of distress expressions, including "help", "help me", "oh my God", "oh", and *"ah"*. In parallel, impact detection compares the audio amplitude with an empirically determined threshold of approximately −45 dBFS, allowing sudden impact sounds to be distinguished from typical indoor background noise. Audio cues alone are not considered sufficient to confirm a fall. Instead, they serve as an initial trigger or as supporting evidence during multimodal decision fusion, consistent with the limitations of audio-only fall detection reported in previous studies [12]. Table 3 summarizes the parameters used by the audio module.

2.6. Rule-based multimodal decision fusion

The outputs of the vision, PIR, and audio modules are combined using a rule-based multimodal fusion strategy that classifies each detected event into one of three priority levels. A rule-based approach was selected because it does not require large, labelled datasets, provides transparent and interpretable decision making, and can be executed efficiently on the Raspberry Pi 4 for real-time operation [20]. Although learning-based fusion offers greater adaptability, its investigation is beyond the scope of this work and is considered future research [15], [16], [21].

The sensing modules operate concurrently using a multi-threaded architecture. Dedicated daemon threads continuously monitor the PIR sensor and audio stream, forwarding trigger events to the main processing loop through a thread-safe queue. This design enables uninterrupted low-power sensing while allowing the camera to be activated only when verification is required.

Vision-based posture analysis provides the primary evidence for confirming a fall, whereas the PIR sensor and audio module serve as event triggers and supporting information. Consequently, isolated motion or audio events, such as those caused by pets or television sounds, cannot generate high-priority alerts without visual confirmation. All threshold values, fusion logic, and timing parameters are fully specified in sections 2.2 to 2.6 and Tables 2 to 4 to support reproducibility, and the full implementation is publicly available (see data availability statement). Table 4 summarizes the rule-based fusion logic, linking PIR/audio triggers and visual confirmation to the resulting priority level and system action.

Table 4. Summarizes the rule-based fusion logic, linking PIR/audio triggers and visual confirmation to the resulting priority level and system action

Priority level	Trigger source	Vision-fall flag	Alert action	Notification
High Confirmed fall	PIR or audio	True	Immediate alert issued	Email + Telegram + Favoriot
Medium Suspected distress	Audio only	False	Reduced-urgency alert	Email + Telegram + Favoriot
Info - no fall confirmed	PIR only	False	Local record only	None

2.7. Alert generation and event logging

Following multimodal decision fusion, the system generates notifications according to the assigned priority level. High-priority events triggered immediate alerts transmitted through email using Python's smtplib library via Gmail simple mail transfer protocol (SMTP) secure sockets layer (SSL), and through Telegram using the hypertext transfer protocol (HTTP) bot application programming interface (API) via the requests library, each containing the event type, timestamp, and a brief description of the detected incident. When visual confirmation is available, a privacy-protected snapshot of the detected posture is also included to help caregivers quickly assess the situation.

Medium-priority events, corresponding to detected audio distress without visual fall confirmation, generate a reduced-urgency alert through Email, Telegram, and Favoriot, informing caregivers of a potential

concern without triggering an emergency response. Info-level events, generated when PIR-only inactivity is detected without supporting visual or audio evidence, are recorded locally without external notification. In all cases, structured event data - including timestamps, trigger sources, priority level, and sensor readings - are logged onto the Favoriot IoT platform to support remote monitoring and historical analysis. Raw visual and audio streams remain on the edge device; only structured event information and, when available, a privacy-protected event snapshot are transmitted for caregiver notification.

3. RESULTS AND DISCUSSION

3.1. Experimental setup

The proposed system was evaluated in a controlled indoor environment designed to simulate a smart-home setting. One volunteer performed ten trials for each of five activity scenarios standing, sitting, kneeling, forward fall, and backward fall resulting in a total of 50 trials. All reported results were obtained using the final calibrated thresholds described in section 2.3. These scenarios are consistent with controlled multimodal fall-detection evaluations reported in the literature [22]. Figure 3 shows the experimental setup integrating the Raspberry Pi 4, Pi Camera, PIR sensor, and USB microphone used for the controlled trials.

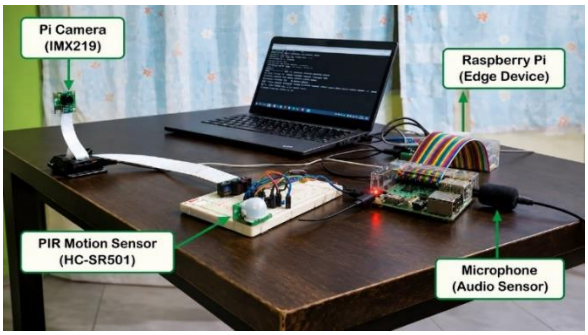
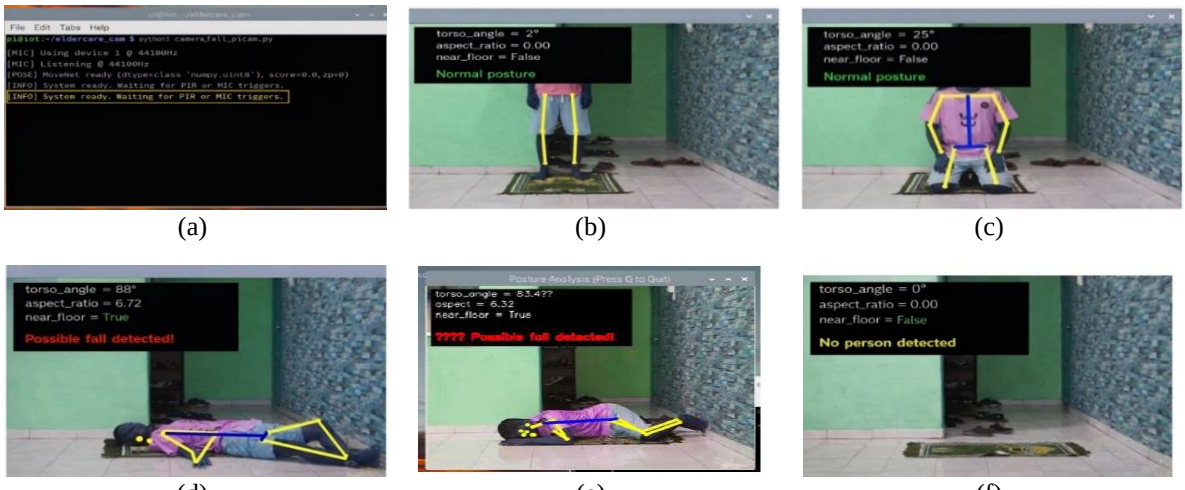


Figure 3. Experimental setup

3.2. System operation verification

The proposed system successfully differentiated between normal activities and fall events. During idle conditions, the camera remained inactive, conserving energy and maintaining privacy. Standing and kneeling postures produced non-fall classifications, whereas forward and backward falls generated high-priority alerts. Figure 4 shows the corresponding outputs for Figures 4(a)-(f).



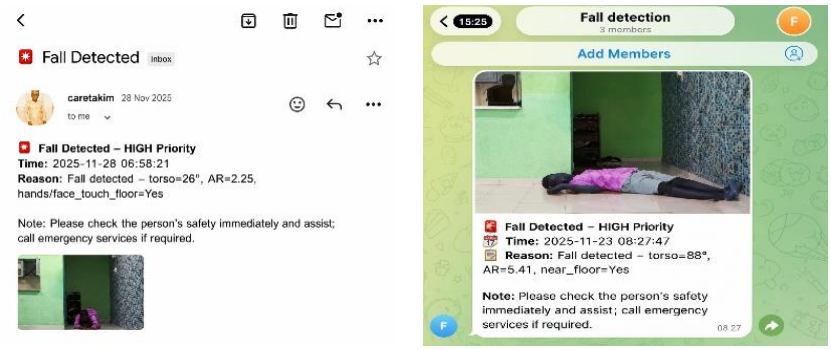
(a) (b) (c)

(d) (e) (f)

Figure 4. Experimental results for different activities and system states: (a) idle, (b) standing, (c) kneeling, (d) backward fall, (e) forward fall, and (f) no person detected

3.3. Alert notification

All confirmed fall events generated alerts through the configured communication channels. Email notifications were successfully delivered for high-priority events, while Telegram messages provided immediate mobile notifications. Favoriot stored event metadata for remote monitoring. No alert delivery failures were observed during the controlled trials. Figure 5 shows the corresponding Figures 5(a) and (b).



(a) (b)

Figure 5. Alert notifications generated by the proposed system: (a) email alert and (b) telegram alert

3.4. Detection performance

Detection performance was evaluated using the confusion matrix in Table 5. Of the 50 trials, the system produced 20 TP, 29 TN, 1 FP, and no FN. Based on the confusion matrix in Table 5, the resulting performance metrics are summarized in Table 6.

Table 5. Confusion matrix of the proposed system

Actual/predicted	Fall	Non-fall
Fall	20 (TP)	0 (FN)
Non-fall	1 (FP)	29 (TN)

Table 6. Detection performance metrics of the developed system

Metric	Value (%)
Accuracy	98.00
Precision	95.24
Recall (sensitivity)	100.00
F1-score	97.56
Specificity	96.67

The system achieved 100% recall with no missed fall events, which is important for reducing delays in emergency assistance during prolonged unattended falls [2], [4]. The only false positive occurred when a participant leaned significantly forward while sitting, which is considered an acceptable trade-off since missed falls may have more serious consequences than false alarms.

The performance is mainly attributed to the rule-based multimodal fusion, where visual posture verification is used for the final high-priority alert decision. Therefore, PIR or audio triggers alone, such as those caused by pets or television sounds, do not generate fall alarms unless the visual fall conditions are also satisfied. However, due to the limited sample size, the results should be considered a controlled proof-of-concept rather than population-level performance. Confidence intervals were not computed, and larger-scale validation is proposed as future work in section 3.7.

3.5. Benchmark comparison

To assess its relative performance, Table 7 compares the proposed system with recent fall-detection approaches [17], [23] using standard detection metrics. The proposed system achieved the highest recall, a safety-critical metric, while maintaining comparable accuracy without continuous camera operation. Neither benchmark uses event-driven camera activation or reports quantitative energy savings, highlighting the practical contribution of the proposed architecture.

Table 7. Benchmark comparison

Metric	Vargas *et al.* [23] (%)	Chang *et al.* [17] (%)	Proposed system (%)
Accuracy	94.80	98.10	98.00
Precision	96.40	-	95.24
Recall	93.10	98.60	100.00
Specificity	96.60	-	96.67

3.6. System responsiveness and energy efficiency

The average end-to-end response time was approximately 9 s, comprising ~1 s for trigger detection, ~5 s for vision verification, and ~3 s for alert dispatch, as illustrated in Figure 6. No significant response-time outliers were observed during the trials. Energy consumption was estimated using (5).

$$E = P \times t \tag{5}$$

where P is power consumption and t is operating time. Manufacturer-rated powers were 5.0/2.7 W (active/idle) for Raspberry Pi 4, 1.2 W for the Pi Camera, 0.0003 W for the PIR sensor, and 0.1 W for the USB microphone. Based on the observed camera duty cycle of ~10 min/h, continuous and event-driven monitoring consumed an estimated 6.30 and 3.38 Wh over one hour, respectively, corresponding to a 46.3% reduction, as summarized in Table 8. The saving is mainly attributed to camera-off-by-default operation and duty-cycled vision processing. Event-triggered edge processing has also demonstrated reduced power consumption in smart-home monitoring [24]. Evaluation was limited to Raspberry Pi 4; other platforms remain for future investigation.



Figure 6. End-to-end response time across the fall-detection stages

Table 8. Estimated hourly energy consumption of continuous and event-driven monitoring

System type	Camera mode	Active power (W)	Idle power (W)	Energy in 1 h (Wh)
Continuous monitoring	Always on	6.3	-	6.30
Event-driven (developed)	Trigger-based	6.3	2.8	3.38

3.7. Discussion and limitations

The results demonstrate reliable detection, practical latency, and substantial energy savings from the event-driven architecture, while the single false positive suggests that temporal consistency analysis could further improve precision. However, the evaluation was limited to one volunteer, simulated falls, a controlled environment, and a small sample without confidence intervals. Practical deployment may be affected by poor lighting, occlusion, room layout, acoustic variation, pets, and multiple residents. These challenges could be addressed through low-light imaging, strategic sensor placement, room-specific visual/audio calibration, and additional sensing nodes. Pet-related triggers could be filtered during visual verification, while multi-resident environments would require coordinated sensing and resident-aware event association. In multi-room deployments, simultaneous sensor events and vision processing may increase Raspberry Pi 4 central processing unit (CPU) and memory usage and detection latency. Scalability could be improved through distributed edge nodes, lightweight vision models, and central gateway coordination to reduce individual-device processing loads. Future work will evaluate these adaptations through larger-scale, multi-environment testing, cross-platform benchmarking, and learning-based multimodal fusion toward broader IoT-enabled AAL applications [25].

3.8. Ethical and user-centred considerations

Informed consent was obtained from the participant, and raw visual and audio data were processed locally without external transmission. For real-world deployment, continuous monitoring may raise concerns about privacy, stigma, and the balance between user autonomy and safety. Configurable privacy settings should allow users to control sensing and notification preferences, while alert cancellation or acknowledgement could reduce unnecessary escalation. Future evaluations should also incorporate user and caregiver feedback to assess acceptability, usability, and trust in home environments.

4. CONCLUSION

This work developed an event-driven, privacy-preserving, edge-based multimodal fall detection system for smart-home elderly monitoring, integrating MoveNet-based posture analysis, PIR motion sensing, and audio distress detection on a Raspberry Pi 4. Across 50 controlled trials, the system achieved 98.00% accuracy and 100% recall, outperforming benchmark systems in recall while reducing hourly energy consumption by 46.3% through camera-off-by-default operation. By activating vision processing only when triggered by low-power sensors, and processing all data locally, the system offers a practical, low-cost pathway for caregivers and families to monitor elderly residents without the privacy trade-offs of always-on cameras. As the evaluation was limited to a single volunteer in a controlled setting, these results should be viewed as proof-of-concept rather than a population-level performance estimate. Real-world multi-environment deployment, larger and more diverse participant groups, and learning-based multimodal fusion remain key directions for future work.

ACKNOWLEDGMENTS

The authors would like to thank their supervisors and colleagues for their guidance and constructive feedback during the development of this project. The authors also acknowledge the technical support provided during the system implementation and testing.

FUNDING INFORMATION

The authors declare that no funding was received for this research.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sufyan Yazid Abubakar	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Rozeha Binti A. Rashid		✓				✓		✓		✓	✓	✓	✓	
Zakarie Abdi Mohamud		✓	✓	✓			✓		✓	✓	✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors confirm that there are no financial, professional, or personal relationships that could have influenced the conduct or reporting of this study.

DATA AVAILABILITY

The source code is publicly available at https://github.com/centbichi/event-driven-fall-detection. Experimental data is available from the corresponding author upon reasonable request.

REFERENCES

[1] E. Tîrziu, A.-M. Vasilevschi, A. Alexandru, and E. Tudora, "Real-Time Fall Monitoring for Seniors via YOLO and Voice Interaction," *Future Internet*, vol. 17, no. 8, no. 324, 2025, doi: 10.3390/fi17080324.
[2] 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, p. 105626, 2022, doi: 10.1016/j.compbiomed.2022.105626.
[3] R. Analia, A. Forster, S.-Q. Xie, and Z. Zhang, "Privacy-Preserving Approach for Early Detection of Long-Lie Incidents: A Pilot Study with Healthy Subjects," *Sensors*, vol. 25, no. 12, p. 3836, Jun. 2025, doi: 10.3390/s25123836.

[4] Y. Wang and T. Deng, "Enhancing elderly care: Efficient and reliable real-time fall detection algorithm," *DIGITAL HEALTH*, vol. 10, pp. 1–11, 2024, doi: 10.1177/20552076241233690.
[5] B. Zieni, M. A. Ritchie, A. M. Mandalari, and F. Boem, "An Interdisciplinary Overview on Ambient Assisted Living Systems for Health Monitoring at Home: Trade-Offs and Challenges," *Sensors*, vol. 25, no. 3, p. 853, Jan. 2025, doi: 10.3390/s25030853.
[6] S.-A. Moraru et al., "Using IoT Assistive Technologies for Older People Non-Invasive Monitoring and Living Support in Their Homes," *Int. J. Environ. Res. Public Health*, vol. 19, no. 10, p. 5890, 2022, doi: 10.3390/ijerph19105890.
[7] M. Dutt, A. Gupta, M. Goodwin, and C. W. Omlin, "An Interpretable Modular Deep Learning Framework for Video-Based Fall Detection," *Applied Sciences*, vol. 14, no. 11, p. 4722, May 2024, doi: 10.3390/app14114722.
[8] A. Ramachandran and A. Karuppiah, "A Survey on Recent Advances in Wearable Fall Detection Systems," *BioMed Research International*, vol. 2020, no. 1, Jan. 2020, doi: 10.1155/2020/2167160.
[9] N. Biancone, C. Bicchielli, F. Ferri, and P. Grifoni, "Multi-Sensor Fall Detection for Smartphones," *Biomedical Journal of Scientific & Technical Research*, vol. 34, no. 4, pp. 26978–26992, 2021, doi: 10.26717/BJSTR.2021.34.005588.
[10] M. Kepski and B. Kwolek, "Detecting human falls with 3-axis accelerometer and depth sensor," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 770–773, doi: 10.1109/EMBC.2014.6943704.
[11] G. Ben-Sadoun, E. Michel, C. Annweiler, and G. Sacco, "Human Fall Detection Using Passive Infrared Sensors with Low Resolution: A Systematic Review," *Clinical Interventions in Aging*, vol. 17, pp. 35–53, 2022, doi: 10.2147/CIA.S329668.
[12] C. Maher, K. A. Dankiw, B. Singh, S. Bogomolova, and R. G. Curtis, "In-Home Evaluation of the Neo Care Artificial Intelligence Sound-Based Fall Detection System," *Future Internet*, vol. 16, no. 6, p. 197, Jun. 2024, doi: 10.3390/fi16060197.
[13] S. S. Jiao, G. Li, G. Zhang, J. Zhou, and J. Li, "Multimodal fall detection for solitary individuals based on audio-video decision fusion processing," *Heliyon*, vol. 10, no. 8, p. e29596, Apr. 2024, doi: 10.1016/j.heliyon.2024.e29596.
[14] A. Garmendia-Orbegozo, M. A. Anton, and J. D. Nuñez-Gonzalez, "Reduction of Vision-Based Models for Fall Detection," *Sensors*, vol. 24, no. 22, p. 7256, Nov. 2024, doi: 10.3390/s24227256.
[15] C. A. U. Hassan *et al.*, "A Cost-Effective Fall-Detection Framework for the Elderly Using Sensor-Based Technologies," *Sustainability*, vol. 15, no. 5, p. 3982, Feb. 2023, doi: 10.3390/su15053982.
[16] A. Shokrollahi, J. A. Persson, R. Malekian, A. Sarkheyli-Hägele, and F. Karlsson, "Passive Infrared Sensor-Based Occupancy Monitoring in Smart Buildings: A Review of Methodologies and Machine Learning Approaches," *Sensors*, vol. 24, no. 5, p. 1533, Feb. 2024, doi: 10.3390/s24051533.
[17] W.-J. Chang, C.-H. Hsu, and L.-B. Chen, "A Pose Estimation-Based Fall Detection Methodology Using Artificial Intelligence Edge Computing," *IEEE Access*, vol. 9, pp. 129965–129976, 2021, doi: 10.1109/ACCESS.2021.3113824.
[18] J.-M. Guerrero-Rodriguez, M.-A. Cifredo-Chacon, C. Cobos Sánchez, and F. Perez-Peña, "Exploiting the PIR Sensor Analog Behavior as Thermoreceptor: Movement Direction Classification Based on Spiking Neurons," *Sensors*, vol. 23, no. 13, p. 5816, Jun. 2023, doi: 10.3390/s23135816.
[19] R. Naccarelli, S. Casaccia, and G. M. Revel, "The Problem of Monitoring Activities of Older People in Multi-Resident Scenarios: An Innovative and Non-Invasive Measurement System Based on Wearables and PIR Sensors," *Sensors*, vol. 22, no. 9, p. 3472, May 2022, doi: 10.3390/s22093472.
[20] A. Alnattah, M. Jajroudi, S. A. N. Fadafen, M. N. Manzari, and S. Eslami, "Artificial Intelligence in Clinical Decision-Making: A Scoping Review of Rule-Based Systems and Their Applications in Medicine," *Cureus*, vol. 17, no. 8, 2025, doi: 10.7759/cureus.91333.
[21] 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, Art. no. 21537, 2024, doi: 10.1038/s41598-024-71545-6.
[22] L. Martínez-Villaseñor, H. Ponce, J. Brieva, E. Moya-Albor, J. Núñez-Martínez, and C. Peñafort-Asturiano, "UP-Fall Detection Dataset: A Multimodal Approach," *Sensors*, vol. 19, no. 9, p. 1988, Apr. 2019, doi: 10.3390/s19091988.
[23] V. Vargas, P. Ramos, E. A. Orbe, M. Zapata, and K. Valencia-Aragón, "Low-Cost Non-Wearable Fall Detection System Implemented on a Single Board Computer for People in Need of Care," *Sensors*, vol. 24, no. 17, p. 5592, Aug. 2024, doi: 10.3390/s24175592.
[24] M. J. C. S. Reis and C. Serôdio, "Edge AI for Real-Time Anomaly Detection in Smart Homes," *Future Internet*, vol. 17, no. 4, p. 179, Apr. 2025, doi: 10.3390/fi17040179.
[25] S. Salvi, L. Garg, and V. Gurupur, "Stage-Wise IoT Solutions for Alzheimer's Disease: A Systematic Review of Detection, Monitoring, and Assistive Technologies," *Sensors*, vol. 25, no. 17, p. 5252, Aug. 2025, doi: 10.3390/s25175252.

BIOGRAPHIES OF AUTHORS



Sufyan Yazid Abubakar received his B.Tech Ed. degree in Electrical Engineering from Abubakar Tafawa Balewa University (ATBU), Bauchi, Nigeria. He is currently pursuing the M.Eng. degree in Wireless Communication and Networks at Universiti Teknologi Malaysia (UTM). His research interests include IoT, edge computing, AI/ML, 5G wireless communications, and privacy-preserving multimodal systems for elderly fall detection. He can be contacted at email: centbichi@gmail.com.



Rozeha Binti A. Rashid received her B.Sc. degree in Electrical Engineering from the University of Michigan, Ann Arbor, USA, and her M.E.E. and Ph.D. degrees in Telecommunication Engineering from Universiti Teknologi Malaysia (UTM). She is a senior lecturer in the Communication Engineering Program, School of Electrical Engineering, UTM, and head of the Telecommunication Software and System (TeSS) research group. She is an IEEE member with more than 140 publications and extensive research project experience. Her research interests include wireless communications, sensor networks, cognitive radio, IoT, and communication technologies. She can be contacted at email: rozeha@utm.my.



Zakarie Abdi Mohamud received his Bachelor's degree in Telecommunication Engineering from Hormuud University, Somalia, in 2022. He is currently pursuing the Master of Engineering degree in Wireless Communication and Networks at Universiti Teknologi Malaysia (UTM). His research interests include the IoT, automated irrigation systems, antenna systems, 5G and beyond wireless communications, and fiber-optic communication networks. His current research focuses on the development of low-cost IoT-based smart systems that integrate real-time sensing, rule-based decision-making, and edge computing for efficient and sustainable applications. He can be contacted at email: zakizahanuuni@gmail.com.