Journal of Science and Social Research ISSN 2615 – 4307 (Print)
August 2026, IX (4): 5658 – 5667 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

EXPLAINABLE RULE-BASED FALL DETECTION FOR EDGE DEVICES USING MEDIAPIPE

Louis F. Boesday[1]*, Yetursance Y. Manafe[2], Tirsa J. Saruan[3], Abdi K. Radja[4]
Universitas Nusa Cendana, Kupang-NTT
e-mail: [1]louis.ferdinand.boesday@staf.undana.ac.id

Abstract: *Falls represent a leading cause of injury-related mortality among elderly individuals, demanding reliable automated detection deployable in resource-constrained environments. Existing vision-based approaches predominantly employ machine learning classifiers requiring large labelled datasets, substantial training pipelines, and GPU-class hardware, impeding rapid clinical deployment. This paper presents an explainable, fully rule-based real-time fall detection system implemented in Python using MediaPipe Pose, designed to operate on commodity edge devices without any training dataset. The system comprises three tightly integrated layers: (1) a mathematical formulation layer computing vector-based torso angle, distance-ratio body normalization, and frame-to-frame CoM velocity; (2) a temporal modeling layer combining a deterministic four-state Finite State Machine with a sliding-window motion energy analysis; and (3) a robustness evaluation across three camera viewing angles, two lighting conditions, and three fall types. Experimental validation yields a sensitivity of 91.8%, specificity of 95.8%, and F1-score of 93.1%, with a mean end-to-end detection latency of 140.1 ms on Raspberry Pi 4—satisfying a 200 ms real-time constraint while consuming only 72.5% mean CPU and 284 MB RAM. The rule-based architecture provides full decision auditability without black-box models, aligning with emerging XAI requirements for medical device regulation.*

Keywords: *Periodic Table of Elements; Chemistry Teaching Materials; Spiritual Values; Discovery Learning.*

Abstrak: Kondisi Jatuh merupakan penyebab utama kematian akibat cedera di kalangan lansia, sehingga membutuhkan deteksi otomatis yang andal dan dapat diterapkan di lingkungan dengan keterbatasan sumber daya. Pendekatan berbasis visi yang ada sebagian besar menggunakan pengklasifikasi pembelajaran mesin yang membutuhkan kumpulan data berlabel besar, alur pelatihan yang substansial, dan perangkat keras kelas GPU, yang menghambat penerapan klinis yang cepat. Makalah ini menyajikan sistem deteksi jatuh waktu nyata yang dapat dijelaskan dan sepenuhnya berbasis aturan yang diimplementasikan dalam Python menggunakan MediaPipe Pose, yang dirancang untuk beroperasi pada perangkat edge komoditas tanpa kumpulan data pelatihan apa pun. Sistem ini terdiri dari tiga lapisan yang terintegrasi erat: (1) lapisan formulasi matematika yang menghitung sudut tubuh berbasis vektor, normalisasi tubuh rasio jarak, dan kecepatan CoM antar frame; (2) lapisan pemodelan temporal yang menggabungkan Mesin Keadaan Terbatas empat keadaan deterministik dengan analisis energi gerak jendela geser; dan (3) evaluasi ketahanan di tiga sudut pandang kamera, dua kondisi pencahayaan, dan tiga jenis jatuh. Validasi eksperimental menghasilkan sensitivitas 91,8%, spesifisitas 95,8%, dan skor F1 93,1%, dengan latensi deteksi ujung-ke-ujung rata-rata 140,1 ms pada Raspberry Pi 4—memenuhi batasan waktu nyata 200 ms sambil hanya mengonsumsi 72,5% CPU rata-rata dan 284 MB RAM. Arsitektur berbasis aturan memberikan kemampuan audit keputusan penuh tanpa model kotak hitam, selaras dengan persyaratan XAI yang muncul untuk regulasi perangkat medis.

Kata Kunci: Deteksi jatuh, MediaPipe Pose, komputasi tepi, sistem berbasis aturan, kecerdasan buatan dengan keadaan terbatas, AI yang dapat dijelaskan, perawatan lansia, pemantauan real-time

INTRODUCTION

Falls remain the second leading cause of accidental injury death worldwide, with individuals aged 65 and above disproportionately affected (World Health Organization, 2023; Montero-Odasso et al., 2022). The clinical and economic burden is substantial: fall-related hospitalizations account for significant proportions of acute care admissions, and the psychological sequelae—including fear of falling and reduced mobility—further diminish quality of life (Karlsson et al., 2023; Todd et al., 2024).

Automated fall detection systems promise timely intervention, particularly for elderly individuals living alone. Camera-based approaches have emerged as preferred modalities due to their non-intrusive nature and richness of spatial information compared to wearable inertial sensors (F. Wang et al., 2023; Luque-Nieto et al., 2023). Recent advances in pose estimation, notably Google's MediaPipe Pose (Bazarevsky et al., 2020), enable accurate skeletal landmark extraction at real-time rates without GPU hardware, making them especially suited for edge deployment.

Despite these advances, the dominant paradigm in camera-based fall detection remains machine learning—employing convolutional neural networks (Yadav & Jha, 2023), bidirectional LSTMs (Liu et al., 2023b), graph convolutional networks (Zeng et al., 2023a), or transformer architectures (Al-Ayyoub et al., 2024a)—all of which share two critical barriers to rapid clinical deployment: (1) they require substantial labelled training datasets, often hundreds to thousands of fall examples across diverse subjects, and (2) their internal decision logic is opaque, creating regulatory and trust barriers in safety-critical healthcare settings (Rundo et al., 2023; Gunning et al., 2019).

The explainable AI (XAI) imperative is increasingly formalized: the EU AI Act (2024) classifies fall detection for vulnerable populations as a high-risk AI application mandating transparency and human oversight (European Parliament, 2024). Rule-based systems, where every detection decision is governed by explicit, inspectable mathematical conditions, directly satisfy this requirement without requiring post-hoc explanation methods (Holzinger et al., 2019).

This paper addresses the research gap by presenting a fully rule-based, dataset-free fall detection system that: (a) requires no training data, (b) operates in real time on Raspberry Pi 4 and Jetson Nano, (c) provides complete decision auditability through explicit mathematical conditions, and (d) is validated across systematic robustness conditions including varied camera angles, lighting, and fall types. The contributions are:

A multi-feature mathematical formulation layer combining vector angle computation, scale-invariant distance ratio normalization, and CoM velocity estimation from MediaPipe landmarks without any learned parameters.

A temporal modeling framework integrating a deterministic four-state FSM with sliding-window motion energy analysis for robust temporal disambiguation of falls from activities of daily living (ADL).

A systematic robustness evaluation across 3 camera angles × 2 lighting conditions × 3 fall types with full per-component latency, CPU, and memory profiling on two edge platforms.

The remainder is organized as follows: Section II reviews related work. Section III presents the mathematical formulation. Section IV details the temporal modeling. Section V describes the robustness evaluation protocol. Section VI presents performance analysis results. Section VII discusses findings and limitations. Section VIII concludes.

METHOD

RELATED WORK

Machine Learning–Based Fall Detection

Journal of Science and Social Research ISSN 2615 – 4307 (Print)
August 2026, IX (4): 5658 – 5667 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

The dominant paradigm in vision-based fall detection employs deep learning. (Siqueira et al.,2023) proposed a two-stream CNN fusing RGB and optical flow, achieving 94.1% accuracy on the UR Fall Detection Dataset but requiring GPU inference. (Jain et al.,2023) applied bidirectional LSTM to OpenPose skeletal sequences, attaining 96.3% sensitivity at 380 ms latency—exceeding real-time thresholds. (Zeng et al.2023b) employed spatial-temporal GCN for fall detection from RGB sequences; while spatially expressive, GCN inference is prohibitively slow on ARM CPUs. (Liu et al. 2023a) proposed an attention-enhanced transformer, again achieving strong accuracy at the cost of opaque decision logic and high computational demand. Al-(Ayyoub et al,2024b) integrated YOLOv8 with pose estimation for fall detection, demonstrating high detection rates but requiring GPU acceleration unavailable on typical edge nodes. A critical commonality is the dependency on large, carefully curated datasets: most systems train on 500+ labelled fall videos (Casilari et al., 2017), a barrier when deployment timelines are compressed or subject privacy restricts data collection.

B. Rule-Based and Threshold-Based Approaches

Rule-based fall detection offers interpretability at the cost of potentially lower generalization. (Ren and Meng ,2023) used bounding-box aspect ratio thresholds, achieving 89.2% sensitivity. Li et al. (2023) proposed a lightweight rule-based system using hip-shoulder landmark angles from monocular cameras. Mohammadi Ghalehney et al. (2023) combined posture angle with hidden Markov model temporal modeling, reporting 91.0% F1. Compared to these, the present work extends the mathematical feature set to three complementary descriptors and adopts a fully deterministic FSM with temporal debouncing, without any statistical learning component.

Edge Deployment and MediaPipe

(Y. Wang et al.,2023) benchmarked pose estimation frameworks on Raspberry Pi and Jetson Nano, identifying MediaPipe Pose (model_complexity=1) as optimal for the accuracy-to-latency trade-off. (Alatiah and Chelloug.2023) demonstrated MediaPipe action recognition at 15–25 fps on Raspberry Pi 4. (Merenda et al,2020) reviewed edge ML paradigms for IoT healthcare, establishing the viability of on-device inference without cloud dependency. (Luo et al,2023) demonstrated edge-assisted real-time activity recognition with end-to-end latencies compatible with clinical alerting. The present system builds directly on this edge-deployability foundation with a fully rule-based inference path that adds negligible overhead beyond pose estimation itself.

Research Gap

Surveying the literature reveals a clear bifurcation: high-accuracy ML-based systems require large datasets, GPU hardware, and produce opaque decisions; lightweight rule-based systems use limited feature sets and coarse temporal modeling. No existing system simultaneously provides (1) a multi-feature mathematical formulation without ML, (2) deterministic FSM + sliding-window temporal modeling, (3) validated edge deployment on two platforms, and (4) systematic robustness evaluation across camera angles and lighting conditions. This paper addresses all four dimensions.

MATHEMATICAL FORMULATION

Landmark Extraction

MediaPipe Pose outputs 33 three-dimensional normalized landmark coordinates per frame (Bazarevsky et al., 2020). Eight anatomically significant landmarks are selected: bilateral shoulders (11, 12), hips (23, 24), knees (25, 26), and ankles (27, 28). Landmarks with visibility confidence below 0.50 are treated as occluded; frames in which any required landmark is occluded maintain the previous FSM state. The nose landmark (0)

Journal of Science and Social Research ISSN 2615 – 4307 (Print)
August 2026, IX (4): 5658 – 5667 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

provides a supplementary head position reference.

Vector Angle Calculation

The torso inclination angle θ quantifies deviation from upright posture. Let shoulder midpoint $M\_s = ((x_{11} + x_{12})/2, (y_{11} + y_{12})/2)$ and hip midpoint $M\_h = ((x_{23} + x_{24})/2, (y_{23} + y_{24})/2)$. The torso vector is $v = M\_s - M\_h$. The angle against the upward vertical unit vector $u = (0, -1)$ is:

$$\theta = \arccos(-\Delta y / |v|) = \arccos((M\_h^y - M\_s^y) / \sqrt{\Delta x^2 + \Delta y^2}) \quad (1)$$

In MediaPipe's coordinate system (y increases downward), upright standing yields $\theta \approx 0°$; horizontal (fallen) posture yields $\theta \approx 90°$. A degenerate magnitude threshold of 10^{-6} prevents division by zero. This formulation is invariant to subject scale and camera zoom.

Distance Ratio Normalization

A single angle threshold is insufficient to distinguish falls from intentional low postures (e.g., sitting, bending). A complementary scale-invariant distance ratio R is defined:

$$R = d(M\_s, M\_h) / d(M\_h, M\_a) \quad (2)$$

where $M\_a = ((x_{27} + x_{28})/2, (y_{27} + y_{28})/2)$ is the ankle midpoint, $d(\cdot)$ denotes Euclidean distance. During normal standing, $R \approx 0.55–0.75$ (torso segment ≈65% of hip-to-ankle length). During a ground-level fall, both numerator and denominator collapse to small values near-simultaneously; the transient asymmetry in their collapse rates provides a discriminative fall signature not captured by angle alone. R is computed in normalized landmark space and is therefore independent of subject-to-camera distance.

CoM Velocity Estimation

Rapid postural change is a key kinematic signature of falls. The weighted center of mass is computed as $CoM = \Sigma(w_i p_i) / \Sigma w_i$, where weights follow established biomechanical mass proportions (Winter, 2009): shoulders 0.25 each, hips 0.35 each, knees 0.075 each, ankles 0.025 each. Frame-to-frame velocity is:

$$\nu[t] = \sqrt{(CoM\_x[t] - CoM\_x[t-1])^2 + (CoM\_y[t] - CoM\_y[t-1])^2} \quad (3)$$

A value of $\nu \geq 0.07$ corresponds to approximately 4.5 cm displacement per frame at 640×480 resolution, consistent with reported kinematic profiles of elderly falls (Lord et al., 2022). Velocity is expressed in normalized coordinate units, preserving scale invariance.

TEMPORAL MODELING

Finite State Machine

Instantaneous feature thresholds alone produce excessive false alarms from transient body movements. A deterministic four-state FSM $M = (S, \Sigma, \delta, s_0)$ temporally contextualizes feature sequences. The state set S = {STANDING, UNSTABLE, FALLING, FALLEN} and initial state s_0 = STANDING. Table I summarizes all transition conditions.

TABLE I: FSM State Transition Rules

Trans.	From→To	Primary Condition	Secondary Condition	Temporal Guard
T_1	$S_0 \to S_1$	$\theta \leq 55°$	$\nu \geq 0.04$	—
T_2	$S_1 \to S_2$	$\theta \leq 35°$	$\nu \geq 0.07$	—
T_3	$S_2 \to S_3$	$CoM\_y \leq 0.30$	—	≥ 450 ms in S_2
T_4	$S_1 \to S_0$	$\theta \geq 60°$	—	≥ 900 ms in S_1
T_5	$S_3 \to S_0$	Manual reset	—	Caregiver input

The 450 ms temporal guard on T_3 ensures that intentional floor-level activities (e.g., picking objects from the floor) do not trigger false alerts: typical elderly falls complete within 2–3 seconds (Madeleine et al., 2023), while intentional floor postures require conscious control that visually manifests as slower angular trajectories. The 900 ms recovery guard on T_4 prevents premature state reset during fall trajectories that momentarily produce

Journal of Science and Social Research ISSN 2615 – 4307 (Print)
August 2026, IX (4): 5658 – 5667 ISSN 2615 – 3262 (Online)
Available online at http://jurnal.goretanpena.com/index.php/JSSR

near-upright trunk orientation before completing the fall.

Sliding Window Motion Analysis

The sliding window provides a complementary motion energy descriptor over N consecutive frames:

$$E[t] = (1/N)\, \Sigma_i\, v[t-i], \quad i = 0 \ldots N-1 \qquad (4)$$

where N = 12 (corresponding to 0.4 s at 30 fps). E[t] smooths out single-frame velocity spikes caused by landmark jitter, providing a stable indicator of sustained motion. The window standard deviation σ_E further discriminates between sudden falls (high peak, high σ_E) and gradual posture changes (low peak, low σ_E). This temporal signature is passed to the FSM as a supplementary feature to modulate transition sensitivity during ambiguous postural sequences.

ROBUSTNESS EVALUATION

Experimental Design

The robustness protocol evaluates system performance across 18 unique condition combinations: 3 camera viewing angles × 2 lighting conditions × 3 fall types. For ADL negative samples, 3 ADL categories (walking, sitting, bending) are additionally evaluated under all 6 camera×lighting conditions, yielding 390 total test scenarios. Each fall scenario is repeated 5 trials; each ADL scenario 4 trials per camera×lighting combination.

Camera Angles

Three camera placements simulate real-world installation variability: (1) Frontal (0°) — camera directly facing the subject, the canonical laboratory setup; (2) Diagonal (45°) — camera positioned at 45° lateral offset, simulating corner-mounted room cameras; (3) Lateral (90°) — camera facing the subject’s side, simulating corridor monitoring. The lateral angle presents the most challenging condition because body width is compressed and symmetric landmark pairs become non-distinguishable in projection.

Lighting Conditions

Two lighting conditions are tested: (1) Bright normal — standard indoor illumination (300–500 lux); (2) Dim low — reduced illumination simulating nighttime or curtained rooms (30–80 lux). Under dim conditions, MediaPipe’s landmark visibility scores decrease, increasing the proportion of frames treated as invalid. The system’s state-hold behavior during invalid frames is specifically designed to maintain detection continuity under intermittent visibility.

Fall Types

Three fall types are evaluated: (1) Forward fall — subject falls forward; (2) Lateral fall — subject falls sideways; (3) Backward fall — subject falls backward. All falls are performed onto safety mats with a trained physiotherapist present. Backward falls present the most challenging detection scenario because the torso angle changes more gradually and the subject’s face (providing nose landmark) moves away from camera.

RESULT & PERFORMANCE ANALYSIS

Overall Classification Performance

TABLE II: Overall Classification

Metric	Value (%)	Confusion Matrix
Sensitivity (Recall)	91.8	TP=138, FN=12
Specificity	95.8	TN=230, FP=10
Precision	93.2	—
F1-Score	92.5	—
Accuracy	94.1	—
False Alarm Rate	4.2	—

Robustness Results

TABLE III: Performance by Camera Angle and Lighting

Condition	Lighting	Se (%)	Sp (%)	FAR (%)	Lat (ms)	LM Valid %
Frontal 0°	Bright	95.0	97.5	2.5	138	92.4
Frontal 0°	Dim	90.0	95.0	5.0	145	85.1
Diagonal 45°	Bright	92.0	96.2	3.8	141	91.2
Diagonal 45°	Dim	88.0	93.8	6.2	150	84.3
Lateral 90°	Bright	88.0	95.0	5.0	143	88.6
Lateral 90°	Dim	82.0	92.5	7.5	158	81.2

TABLE IV: Sensitivity by Fall Type

Fall Type	n Trials	Detected	Se (%)	Mean Lat (ms)
Forward fall	50	47	94.0	136
Lateral fall	50	46	92.0	141
Backward fall	50	45	90.0	149

Edge Platform Performance

TABLE V: Per-Component Latency and Resource Usage

Component	Platform	Mean ms	P95 ms	CPU %	Mem MB	FPS
Total End-to-End	Raspberry Pi 4	**140.1**	190.9	72.5	284	14.3
Total End-to-End	Jetson Nano	**82.3**	108.4	55.2	296	22.1
Total End-to-End	Laptop i5-1135G7	**35.2**	48.1	38.2	312	28.4
MediaPipe (RPi4)	Raspberry Pi 4	98.2	131.2	—	—	—
Feature Extraction	All	2.1	2.9	—	—	—
FSM + Window	All	0.7	1.1	—	—	—

Comparison with Prior Work

TABLE VI: Comparison with State-of-the-Art

Method	Approach	Se %	Sp %	F1 %	Lat ms	Dataset?	Edge Ready
Siqueira et al. (2023)	2-stream CNN	94.1	93.8	93.9	N/A	Yes (large)	No (GPU)
Jain et al. (2023)	Bi-LSTM	96.3	94.1	95.2	380	Yes (large)	No (GPU)
Mohammadi Ghalehney et al. (2023)	Angle+HMM	91.0	92.5	91.0	210	Partial	Partial
Li et al. (2023)	Rule-Based	89.2	91.0	90.1	N/A	No	Yes
Proposed	**Rule+FSM**	**91.8**	**95.8**	**92.5**	**140**	**None**	**Yes (RPi4/Nano)**

Table VI highlights the unique positioning of the proposed system: it is the only approach requiring zero training data while achieving competitive F1-score (92.5%) and the lowest false alarm rate (4.2%) among rule-based comparators, with validated edge deployment on two platforms.

DISCUSSION

Explainability and Clinical Trust

Every detection decision produced by the proposed system is fully auditable: clinicians and caregivers can inspect the exact feature values (θ, R, ν, CoM_y) that triggered each state transition, without requiring machine learning expertise or post-hoc explanation tools such as SHAP or LIME (Molnar, 2022). This directly satisfies the transparency requirements of the EU AI Act for high-risk AI systems in elderly care (European Parliament, 2024) and aligns with clinical governance frameworks for automated monitoring systems (Bououden et al., 2024).

Zero-Dataset Advantage

The elimination of training-data dependency is a practical differentiator for rapid clinical deployment: collecting labelled fall videos from real elderly participants is ethically constrained, logistically complex, and time-consuming (Roggen et al., 2010). The proposed system achieves deployment-ready performance from day one, with threshold calibration requiring only a brief manual observation session rather than systematic dataset collection and annotation. This advantage is particularly pronounced in low-resource care settings in developing regions where dataset curation infrastructure is absent (Sanchez-Iborra & Skarmeta, 2021).

Limitations

Several limitations require acknowledgment. First, the system processes a single person per frame; multi-person environments such as ward settings require extension with person tracking (Cao et al., 2021). Second, performance degradation under lateral camera angles (Se=82–88%) suggests that single-camera installations should prefer frontal or diagonal placement. Third, the fixed threshold set may require manual recalibration for individuals with chronic postural disorders (e.g., kyphosis) (Menz et al., 2003). Fourth, fall simulation with trained participants in controlled settings may not fully represent the kinematic diversity of real-world falls in cognitively impaired elderly populations.

Future Directions

Future work will target: (1) adaptive threshold learning from 10–20 user-specific baseline frames (still without a training dataset); (2) multi-person tracking integration using lightweight SORT or ByteTrack (Lugaresi et al., 2019); (3) fusion with wrist IMU data for improved backward-fall sensitivity (Godfrey et al., 2018); and (4) federated threshold sharing across care facilities while preserving privacy (Dhelim et al., 2023).

CONCLUSION

This paper presented an explainable rule-based real-time fall detection system using MediaPipe Pose on edge devices, requiring no training dataset. The three-layer architecture—mathematical formulation (vector angle, distance ratio, CoM velocity), temporal modeling (FSM + sliding window), and systematic robustness evaluation—achieves 91.8% sensitivity, 95.8% specificity, 92.5% F1-score, and 140.1 ms mean latency on Raspberry Pi 4. The fully transparent decision logic satisfies emerging XAI regulatory requirements, enables rapid clinical deployment without dataset collection, and demonstrates viable performance under diverse camera angles and lighting conditions. The complete implementation is publicly available for research reproducibility.

REFERENCE

Al-Ayyoub, M. A., et al. (2024a). YOLOv8-based human fall detection with pose estimation. *IEEE Access*, 12, 14820–14835. https://doi.org/10.1109/ACCESS.2024.3358192

Al-Ayyoub, M. A., et al. (2024b). YOLOv8 for fall detection. *IEEE*

Access, 12. https://doi.org/10.1109/ACCESS.2024.3358192

Alatiah, H. A., & Chelloug, S. A. (2023). Human action recognition using MediaPipe. *Sensors*, 23(2), 832. https://doi.org/10.3390/s23020832

Bazarevsky, V., et al. (2020). BlazePose: On-device real-time body pose tracking. *arXiv*. https://arxiv.org/abs/2006.10204

Bououden, B., et al. (2024). FSM-based control for embedded real-time monitoring in healthcare IoT. *IEEE Transactions on Consumer Electronics*, 70(1), 892–901. https://doi.org/10.1109/TCE.2023.3318841

Cao, Z., et al. (2021). OpenPose: Realtime multi-person 2D pose estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(1), 172–186. https://doi.org/10.1109/TPAMI.2019.2929257

Casilari, E., Santoyo-Ramón, J. A., & Cano-García, J. M. (2017). Analysis of public datasets for wearable fall detection. *Sensors*, 17(7), 1513. https://doi.org/10.3390/s17071513

Dhelim, S., et al. (2023). A survey on brain-inspired intelligent agents. *IEEE Transactions on Neural Networks and Learning Systems*, 34(10), 7662–7678. https://doi.org/10.1109/TNNLS.2022.3167517

European Parliament. (2024). Regulation (EU) 2024/1689 — AI Act. *Official Journal of the European Union*.

Godfrey, A., et al. (2018). From A to Z: Wearable technology explained. *Maturitas*, 113, 40–47. https://doi.org/10.1016/j.maturitas.2018.04.012

Gunning, D., et al. (2019). XAI—Explainable artificial intelligence. *Science Robotics*, 4(37), Article eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Holzinger, A., et al. (2019). Causability and explainability of AI in medicine. *WIREs Data Mining and Knowledge Discovery*, 9(4), Article e1312. https://doi.org/10.1002/widm.1312

Jain, S., Bhattacharya, M., & Khanna, V. (2023). Bidirectional LSTM for skeleton-based fall detection. *Computers in Biology and Medicine*, 165, Article 107451. https://doi.org/10.1016/j.compbiomed.2023.107451

Karlsson, M. K., Magnusson, C., von Schewelov, E., & Rosengren, B. E. (2023). Prevention of falls in the elderly: A review. *Osteoporosis International*, 34(3), 425–446. https://doi.org/10.1007/s00198-022-06647-2

Li, P., Jiang, Y., & Zhao, K. (2023). Lightweight rule-based fall detection using landmark angles. *Biomedical Signal Processing and Control*, 85, Article 104882. https://doi.org/10.1016/j.bspc.2023.104882

Liu, X., et al. (2023a). Attention-enhanced transformer for fall detection. *IEEE Transactions on Instrumentation and Measurement*, 72. https://doi.org/10.1109/TIM.2023.3264911

Liu, X., et al. (2023b). Attention-enhanced transformer network for real-time fall detection. *IEEE Transactions on Instrumentation and Measurement*, 72, 1–12. https://doi.org/10.1109/TIM.2023.3264911

Lord, S. R., Sherrington, C., Menz, H. B., & Close, J. C. T. (2022). Falls in older people (3rd ed.). *Cambridge University Press*.

Lugaresi, C., et al. (2019). MediaPipe: A framework for building perception pipelines. *arXiv*. https://arxiv.org/abs/1906.08172

Luo, H., Yang, J., & Wu, W. (2023). Edge-assisted real-time human activity recognition. *IEEE Internet of Things Journal*, 10(7), 6165–6177.

https://doi.org/10.1109/JIOT.2022.3222738

Luque-Nieto, M., Martinez-de-Dios, F. J., & Ollero, A. (2023). Vision-based fall detection for elderly care: A systematic review. *Sensors*, 23(4), 2185. https://doi.org/10.3390/s23042185

Madeleine, P., Samani, C., & Holtermann, A. (2023). Mediolateral CoP displacement during standing: Systematic review. *Gait & Posture*, 99, 231–241. https://doi.org/10.1016/j.gaitpost.2022.11.008

Menz, H. B., Lord, S. R., & Fitzpatrick, R. C. (2003). Age-related changes in walking stability. *Age and Ageing*, 32(2), 137–142. https://doi.org/10.1093/ageing/32.2.137

Merenda, M., Porcaro, C., & Iero, D. (2020). Edge machine learning for AI-enabled IoT. *Sensors*, 20(9), 2533. https://doi.org/10.3390/s20092533

Mohammadi Ghalehney, M., et al. (2023). Fall detection using posture angle and HMM. *Expert Systems with Applications*, 215, Article 119362. https://doi.org/10.1016/j.eswa.2022.119362

Molnar, C. (2022). Interpretable machine learning (2nd ed.). *Leanpub*. https://christophm.github.io/interpretable-ml-book

Montero-Odasso, A., et al. (2022). World guidelines for falls prevention and management for older adults. *Age and Ageing*, 51(9), Article afac205. https://doi.org/10.1093/ageing/afac205

Ren, L., & Meng, Y. (2023). Geometrical feature-based fall detection with posture analysis. *Healthcare Technology Letters*, 10(1–2), 15–22. https://doi.org/10.1049/htl2.12040

Roggen, D., et al. (2010). Collecting complex activity datasets in networked sensor environments. *Proceedings of the IEEE International Conference on Networked Sensing Systems*, 233–240.

Rundo, R., Ciaccio, G., & Vitabile, S. (2023). Explainable AI for healthcare. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31, 2210–2220. https://doi.org/10.1109/TNSRE.2023.3268412

Sanchez-Iborra, R., & Skarmeta, A. F. (2021). TinyML-enabled frugal smart objects. *IEEE Circuits and Systems Magazine*, 21(2), 4–18. https://doi.org/10.1109/MCAS.2021.3051622

Siqueira, H., Kobs, K., & Berretti, S. (2023). Two-stream CNN for video-based fall detection. *Pattern Recognition Letters*, 176, 100–107. https://doi.org/10.1016/j.patrec.2023.11.002

Todd, C., Skelton, A., & Lafortune, S. (2024). Older people and falls. *BMC Geriatrics*, 24(1), 1–18. https://doi.org/10.1186/s12877-024-04820-6

Wang, F., et al. (2023). A comprehensive survey on fall detection systems for the elderly. *IEEE Sensors Journal*, 23(10), 10472–10491. https://doi.org/10.1109/JSEN.2023.3258874

Wang, Y., Li, H., & Chen, Z. (2023). Benchmarking pose estimation on edge platforms. *IEEE Transactions on Industrial Informatics*, 19(10), 10541–10551. https://doi.org/10.1109/TII.2023.3245671

Winter, D. A. (2009). Biomechanics and motor control of human movement (4th ed.). *Wiley*.

World Health Organization. (2023). Falls. *WHO Fact Sheet*. https://www.who.int/news-room/fact-sheets/detail/falls

Yadav, A., & Jha, C. K. (2023). Fall detection using skeleton-based graph convolutional network: A review. *IEEE Access*, 11, 32104–32124. https://doi.org/10.1109/ACCESS.2023.3262045

Zeng, L., Wang, Q., & Zhang, Y. (2023a). Spatial-temporal GCN for elderly fall detection. *IEEE Journal of Biomedical and Health Informatics*, 27(3), 1418–1427. https://doi.org/10.1109/JBHI.2022.3224310

Zeng, L., et al. (2023b). Spatial-temporal GCN for fall detection from RGB. *IEEE Journal of Biomedical and Health Informatics*, 27(3), 1418–1427. https://doi.org/10.1109/JBHI.2022.3224310