JOURNAL OF SCIENCE
OF LAC HONG UNIVERSITY
JSLHU

ISSN: 2525 - 2186
Tạp chí Khoa học Lạc Hồng, **2026**, 26, 018-024

PHÁT HIỆN TẾ NGÃ ĐỘT PHÁ BẰNG CÔNG NGHỆ THỊ GIÁC MÁY TÍNH

Đỗ Trí Nhựt[1,2*], Lê Thị Thủy[3]
[1]*Trường Đại học Công nghệ thông tin, Tp. Hồ Chí Minh, Việt Nam*
[2]*Đại học Quốc gia Tp. Hồ Chí Minh, Tp. Hồ Chí Minh, Việt Nam*
[3]*Trường Đại học Thủ Dầu Một, Tp. Hồ Chí Minh, Việt Nam*
*Tác giả liên hệ: *trinhutdo@gmail.com/trinhutdo@uit.edu.vn*

THÔNG TIN BÀI BÁO

Ngày nhận:	02/05/2025
Ngày hoàn thiện:	05/08/2025
Ngày chấp nhận:	07/08/2025
Ngày đăng:	15/03/2026

TỪ KHÓA

Phát hiện té ngã;
Thị giác máy tính;
Phân tích điểm đặc trưng;
Giám sát thời gian thực;
Tích hợp học máy.

TÓM TẮT

Khi dân số già đi, nhu cầu về các hệ thống giám sát an toàn hiệu quả ngày càng tăng, đặc biệt là đối với các nhóm dễ bị tổn thương như người cao tuổi. Bài báo này giới thiệu một hệ thống phát hiện té ngã mới sử dụng các kỹ thuật thị giác máy tính tiên tiến, chẳng hạn như phân tích điểm đặc trưng và mô hình GRU nhẹ, để cung cấp khả năng giám sát và can thiệp theo thời gian thực trong khi té ngã, với thời gian xử lý khung hình từ 0,02 đến 0,11 giây và thời gian phản hồi luôn dưới 0,09 giây. Hệ thống được đề xuất đạt độ chính xác 95% trong việc phân biệt các hoạt động bình thường với té ngã, giúp giảm đáng kể các kết quả dương tính giả thông qua thuật toán mạnh mẽ và xử lý trước cẩn thận. Bằng cách tích hợp các hệ thống thông minh vào môi trường hàng ngày, phương pháp tiếp cận của chúng tôi chứng minh tính hiệu quả của các phương pháp trích xuất đặc trưng và cho thấy tiềm năng đáng kể cho các ứng dụng chăm sóc sức khỏe. Những phát hiện này nhấn mạnh khả năng nâng cao sự an toàn và hạnh phúc của hệ thống, gợi ý những cải tiến trong tương lai thông qua các mô hình học máy có thể tối ưu hóa khả năng phát hiện theo thời gian. Nghiên cứu này hứa hẹn sẽ cách mạng hóa việc giám sát chăm sóc sức khỏe tự động, thúc đẩy sự độc lập và chất lượng cuộc sống tốt hơn cho những cá nhân có nguy cơ. Chúng tôi hy vọng công trình của mình sẽ truyền cảm hứng cho những tiến bộ hơn nữa trong lĩnh vực quan trọng này, góp phần tạo nên bộ giải pháp an toàn toàn diện cho nhóm dân số già trên toàn thế giới.

A COMPUTER VISION APPROACH FOR INNOVATING FALL DETECTION

Tri Nhut Do[1,2*], Thi Thuy Le[3]
[1]*University of Information Technology, Ho Chi Minh City, Vietnam*
[2]*Vietnam University in HCMC, Ho Chi Minh City, Vietnam*
[3]*Thu Dau Mot University, Ho Chi Minh City, Vietnam*
*Corresponding author: *trinhutdo@gmail.com/trinhutdo@uit.edu.vn*

ARTICLE INFO

Received:	May 2nd, 2025
Revised:	Aug 5th, 2025
Accepted:	Aug 7th, 2025
Published:	Mar 15th, 2026

KEYWORDS

Fall Detection;
Computer Vision;
Feature Point Analysis;
Real-time Monitoring;
Machine Learning Integration.

ABSTRACT

As the population ages, the need for effective safety monitoring systems increases, especially for vulnerable groups such as the elderly. This paper introduces a novel fall detection system that utilizes advanced computer vision techniques, including feature point decomposition and a lightweight GRU model, to deliver real-time monitoring and interventions during falls. The system achieves frame processing times between 0.02 and 0.11 seconds, with response times consistently under 0.09 seconds. The proposed system achieves 95% accuracy in distinguishing between normal activities and falls, significantly reducing false positives through a robust algorithm and careful preprocessing. By integrating intelligent systems into everyday environments, our approach demonstrates the effectiveness of feature extraction methods and reveals significant potential for healthcare applications. These findings underscore the system's ability to enhance safety and well-being, suggesting future improvements through machine learning models that can optimize detection over time. This research promises to revolutionize automated healthcare monitoring, fostering greater independence and quality of life for at-risk individuals. We hope our work inspires further advancements in this vital area, contributing to a comprehensive suite of safety solutions for aging populations worldwide.

Doi: https://doi.org/10.61591/jslhu.26.776
Available online at: https://lhj.vn

1. INTRODUCTION

The global population is undergoing a profound demographic shift, with a growing proportion of older adults who are increasingly susceptible to falls, a leading cause of injury-related mortality and disability. The World Health Organization reports approximately 646,000 annual deaths from falls, predominantly among the elderly, underscoring the urgent need for effective fall prevention and detection systems to enhance safety and support independent living [1-5]. In Vietnam, an estimated 1.5–1.9 million older adults experience falls annually, with 5% requiring hospitalization for injuries. At the Department of Geriatrics and Palliative Care, University of Medicine and Pharmacy Hospital in Ho Chi Minh City, around 17% of monthly admissions stem from fall-related events [6-8], highlighting the local severity of this issue.

Recent technological advancements have spurred the development of fall detection systems, broadly classified into wearable and ambient solutions. Wearable devices, such as accelerometers and gyroscopes, enable continuous motion tracking but are often hindered by user discomfort and limited battery life [9-12]. Conversely, ambient systems leverage environmental sensors for unobtrusive monitoring yet struggle with accuracy due to variable indoor conditions and frequent false positives, where routine movements are mistaken for falls [13]. These limitations necessitate innovative approaches to improve reliability and user acceptance.

To overcome these challenges, computer vision and deep learning have emerged as promising tools for enhancing fall detection. For instance, Sykes used pose estimation to focus on human joint movements relative to environmental features (e.g., proximity to furniture or walls), reducing false positives by contextualizing motion patterns [14]. Building on this, our proposed system introduces a novel approach using landmark identification to contextualize user surroundings, significantly reducing false positives. Recent deep learning advancements, such as convolutional neural networks employed by Ge et al. to distinguish falls from non-falls in controlled settings [15], further bolster the potential for precise movement analysis. This research aims to advance fall detection by integrating these cutting-edge techniques, offering a robust solution to safeguard vulnerable populations.

This paper introduces an innovative fall detection system that integrates landmark identification with deep learning to improve accuracy and reliability in identifying fall events. Leveraging advanced computer vision techniques, including MediaPipe for spatial keypoint detection [16], our system dynamically tracks user movements to differentiate falls from routine activities with high precision. By coupling these landmarks with a deep learning-based prediction model, the approach significantly reduces false positives, ensuring timely and dependable responses. Extensive experiments demonstrate the system's robustness across varied scenarios, underscoring its potential to expedite interventions and enhance safety and independence for older adults. This work advances eldercare technology and lays a foundation for further innovations in smart health monitoring. Moreover, the system is designed for seamless integration with IoT frameworks, enabling user-friendly interfaces that optimize satisfaction for both caregivers and individuals under supervision.

The rest of this paper is organized as follows: Section 2 describes the design of the proposed system. Section 3 presents some experimental results. Section 4 concludes the paper with future directions.

2. DESIGN OF THE PROPOSED SYSTEM

The proposed method develops a system for human body motion recognition and analysis by integrating image processing, deep learning, and the MediaPipe tool. As illustrated in Figure 3, the system utilizes a camera to capture human body frames, which are processed by MediaPipe [17] to extract 33 key points. These data points are then input into a lightweight GRU neural network [18] for behavior recognition and motion analysis. The model is designed for deployment on embedded devices, enabling applications in health monitoring and safety warning.

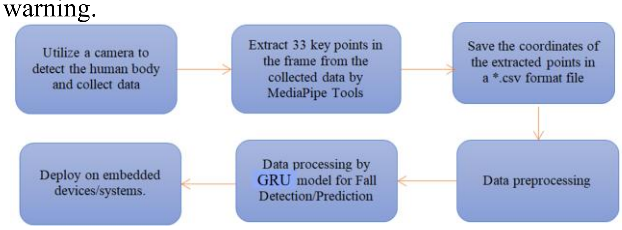

Figure 1. *Human body motion recognition system diagram with MediaPipe tools and GRU model*

2.1 Design by Mediapipe and MediaPipe Pose

MediaPipe is an open-source, cross-platform framework developed by Google since 2010 for real-time multimedia processing applications. It supports machine learning tasks involving images, videos, audio, and sensor data, providing effective tools for face recognition, body pose estimation, gesture recognition, and object detection.

Key features include GPU acceleration for enhanced processing speed and parallel processing capabilities that enable simultaneous handling of multiple video streams or computer vision models. Moreover, MediaPipe seamlessly integrates with OpenCV and TensorFlow, simplifying the incorporation of pre-trained or custom models.

MediaPipe Pose [19] is a powerful tool for human body pose recognition and tracking, utilizing deep learning with Convolutional Neural Networks (CNN) to identify 33 keypoints on the body from images or videos. These landmarks—located at critical points like the head, shoulders, and knees—enable real-time movement tracking and pose estimation. This capability has diverse applications, including rehabilitation monitoring, athlete technique analysis, and fall detection for timely alerts. Figure 2 demonstrates the placement of these landmarks, highlighting how MediaPipe Pose effectively analyzes human body movement.

In fall detection, recognizing and tracking human body motion is crucial. MediaPipe Pose effectively identifies and tracks 33 key landmarks on the body, including the head, shoulders, and limbs, enabling real-time analysis of movements and behavior changes.

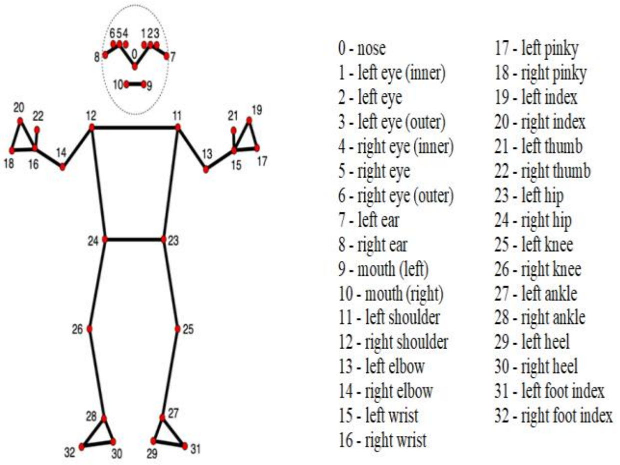

Figure 2. *MediaPipe-based body landmarks [19]*

When a fall occurs, these landmarks—especially around the knees, ankles, arms, and shoulders—shift significantly, signaling a loss of balance. MediaPipe Pose provides data on these changes, aiding machine learning models in recognizing abnormal behaviors like falls.

An effective fall detection system not only identifies falls but also predicts warning signs. Monitoring landmark changes over time can detect unusual patterns, such as sudden posture shifts or uncoordinated movements, that may precede a fall.

The Gated Recurrent Unit (GRU), introduced by Cho et al. (2014) [21], is a simplified variant of the Long Short-Term Memory (LSTM) model, featuring just two main gates: the reset gate and the update gate. This streamlined structure reduces the number of parameters compared to LSTMs, enabling faster training and forecasting, particularly with small datasets. GRUs maintain high performance without excessive computational demands, minimizing the risk of overfitting, making them ideal for problems requiring rapid training on simpler datasets.

Through detailed analysis, we observe that GRUs simplify the complex gating mechanisms of LSTMs while achieving comparable performance in capturing temporal dependencies. A GRU unit consists of two gates—the update and reset gates—that regulate information flow by balancing memory retention with the incorporation of new input. This architecture reduces computational complexity relative to LSTMs, which employ three gates—the input, forget, and output gates—yet remains effective in modeling long-term dependencies. As a result, GRUs are particularly well-suited for processing sequential data in fall detection applications, such as analyzing temporal sequences from wearable sensors or video frames. Additionally, hyperparameter tuning is crucial for optimizing GRU performance in fall detection tasks, where accuracy, processing speed, and robustness to noise are critical. Finally, employing cross-validation ensures a robust evaluation of GRU models, especially given that datasets like Le2i are relatively small and imbalanced, with more fall instances than non-fall ones.

In this section, we have developed an innovative fall detection method using computer vision technology that leverages the Mediapipe tool for motion feature extraction and employs a GRU deep learning model for classification and prediction. This system enhances accuracy over traditional methods and can be easily deployed on surveillance cameras or mobile applications. With its significant advantages, this solution aims to protect the health of the elderly, support smart healthcare, and improve quality of life.

2.2 Dataset

In this section, we describe in detail how the dataset is formed and how the steps involved in training the system are included.

To enhance model flexibility and generalizability, we curated a diverse training dataset by integrating our own data with existing sources, reflecting various body sizes and genders. The dataset includes individuals with different body shapes to ensure accurate performance in real-world scenarios and incorporates data from official datasets such as Le2i and the Multiple Cameras Fall Dataset, which provide body feature coordinates across various poses. Of which, the Le2i Fall Detection Dataset is a robust resource for vision-based fall detection, offering 191–250 videos across four realistic indoor scenes with diverse lighting and activity types. While demographic details are sparse, it likely focuses on simulating elderly fall scenarios with a small participant pool. Its single-camera, low-resolution setup and partial annotations make it ideal for testing real-time, resource-efficient algorithms, but limit its scope for diverse or multimodal studies. For a more tailored analysis or comparison with a specific fall detection system, please provide details about the proposed system or desired metrics.

To maintain input data consistency, we standardized video lengths in the training set to 50 consecutive frames, or approximately 1.5 seconds, focusing on the fall phase. This normalization removes extraneous frames before and after the fall, ensuring a balanced representation of both "fall" and "no fall" classes and reducing noise. Extracting these 50 frames enhances our system's understanding of body kinematics during falls, improving classification accuracy and optimizing model performance. This methodology is detailed in Figure 3.

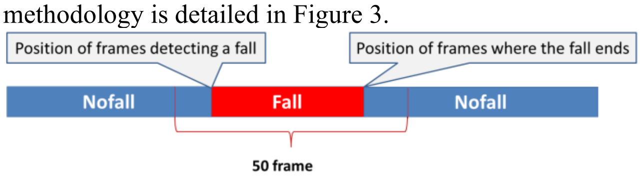

Figure 3. *The sampling method for falls*

Figure 3 illustrates that the fall analysis requires a frame window overlapping by 5 to 10 frames before and after the event, enhancing the accuracy of the fall detection model by capturing the transitional phases and minimizing the risk of missing critical data.

The selection of 50 frames is designed to facilitate real-time fall detection and ensure timely incident recognition. Among these, 30 central frames are chosen to match the standard frame rate perceived by the human eye, ensuring accurate temporal processing. Additionally, 10 frames before and after the central set are included to maintain continuous, real-time processing and improve the system's responsiveness.

As we know, the Mediapipe Pose tool employs a convolutional neural network (CNN) to identify 33 feature points on the human body, each represented by four values: three spatial coordinates (x, y, z) and confidence (visibility). This results in a feature vector of 132 elements (33 points by four values) per video frame. Using a dataset of 516 videos, each with 50 frames labeled as "fall" or "no fall," we generate an input dataset of 25,800 samples with 132 features. Converting the data into vectors enhances the deep learning model's ability to analyze and classify actions, ultimately improving fall prediction accuracy. Finally, the dataset labeled as "fall" is shown in Figure 4, and labeled as "no fall" is shown in Figure 5.

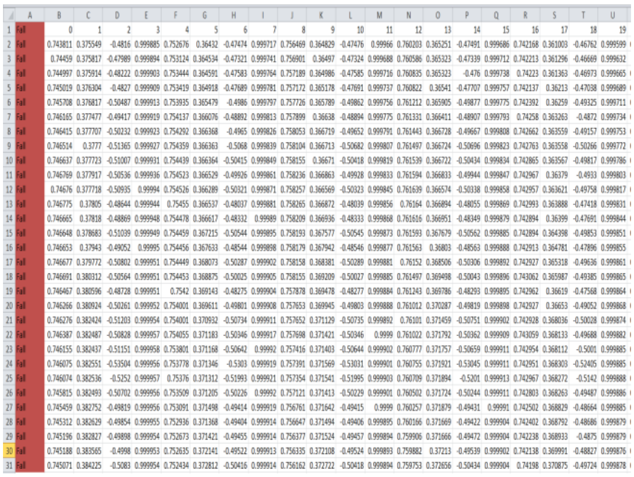

Figure 4. *The dataset is labeled as "fall"*

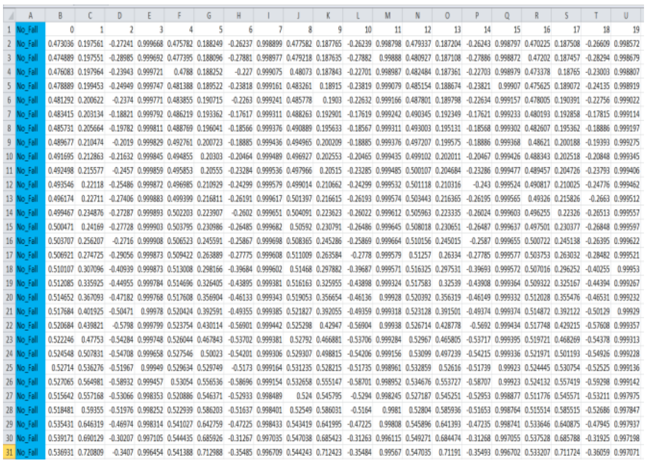

Figure 5. *The dataset is labeled as "no-fall"*

2.3 Training

The training dataset was divided into two parts, with 70% of the dataset for training and 30% for validation. The training input images were preprocessed, and skeletal features were extracted to show the fall prediction metrics.

In this section, the lightweight deep learning GRU (Gated Recurrent Unit) model is implemented for fall detection. The model was trained on the same dataset with identical epochs and optimization parameters to ensure a fair evaluation. Their performance was assessed using a test dataset and calculated metrics, including True Positives, True Negatives, False Positives, and False Negatives, along with accuracy, F1 Score, precision, and recall.

The proposed system's training pipeline comprises three key processes, each supported by a dedicated module: (1) *made_Data.py* preprocesses video data, extracting features to prepare for model training; (2) *train_GRU.py* optimizes GRU models using preprocessed data, enhancing fall detection accuracy; and (3) *inference.py* deploys the trained model for real-time fall prediction, enabling seamless integration into monitoring systems.

2.4 Validating

To rigorously assess the performance of the trained system, quantitative metrics such as precision, recall, and F1-score are indispensable. These metrics are derived from the model's confusion matrix, a crucial representation of its classification accuracy.

For our specific task, we categorize instances as either "fall" (positive, P) or "no fall" (negative, N). Crucially, we define true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) based on the unnormalized confusion matrix as follows:

***Table 1**. General confusion matrix.*

		Predicted Class	
		Possitive	Negative
Actual Class	Possitive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

- **Precision:** a critical metric in evaluating the performance of image classifiers, quantifies the accuracy of a model's positive predictions. A high Precision value, approaching 1, signifies a strong ability to identify only relevant instances as positive. This ideal is achieved when all predicted positives are positive (i.e., True Positives, TP, equal the sum of True Positives and False Positives, TP + FP). Crucially, this implies a zero false positive rate (FP=0). Conversely, as the number of false positives (FP) increases, the denominator (TP + FP) becomes larger than the numerator (TP), and Precision correspondingly decreases. The Precision is formally calculated as the ratio of True Positives to the sum of True Positives and False Positives: **Precision = TP / (TP + FP).**

- **Recall:** a critical metric in evaluating classification models, quantifies the ability of a model to correctly identify all instances of a class. A recall score approaching 1 signifies excellent performance, indicating that the model correctly identifies virtually all instances belonging to the class. Recall attains a value of 1 only when the true positive (TP) count equals the sum of true positives and false negatives (TP + FN), implying a perfect identification of all relevant instances (FN = 0). Conversely, increasing the number of false negatives (FN) directly results in a declining recall score. This index, crucial for assessing the model's generalization capacity to accurately identify relevant instances in new, unseen data, is formally defined by the ratio of true positives to the sum of true positives and false negatives: **Recall = TP / (TP + FN).**

- **F1-score:** Precision and recall are crucial metrics for evaluating classification performance. Ideally, both should be high. However, a pursuit of maximal recall often necessitates a trade-off with precision, and vice versa. Increasing recall to the detriment of precision, or improving precision at the expense of recall, are common pitfalls. Consequently, the F1-score, calculated as a harmonic mean of precision and recall, provides a balanced measure of performance, mitigating the impact of this inherent trade-off. The F1-score is calculated by formula (1).

$$\mathbf{F1-Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \tag{1}$$

- **Acc** ıating computer vision models, quantifies the proportion of correctly classified test instances. It is calculated by dividing the total number of correctly classified samples by the total number of samples in the validating dataset, as in formula (2).

$$\mathbf{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{2}$$

The training results of the system were obtained after 80 iteration cycles (epochs) on the validation dataset. The confusion matrix is illustrated as in Figure 6.

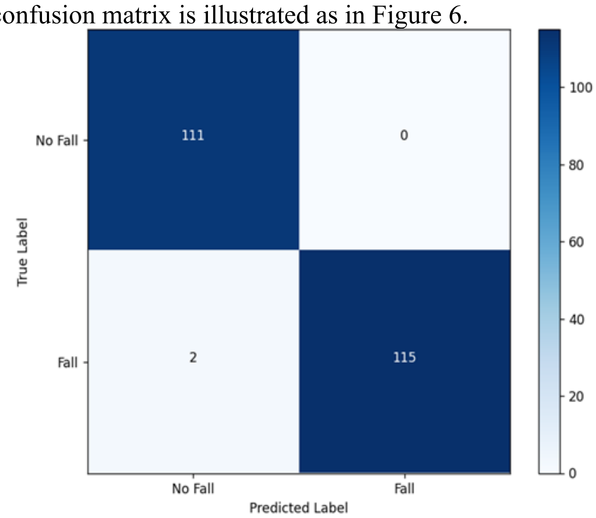

Figure 6. *The confusion matrix*

Based on the data obtained from the training process and represented by the confusion matrix, the quantitative metrics such as precision, recall, and F1-score are calculated and listed as in Table 2.

Table 2*. The quantitative metrics.*

Accuracy	Precision	Recall	F1-Score
99.12%	1.00	0.98	0.99

To enhance academic credibility, the LSTM model was also employed for training and validation. Consequently, quantitative metrics such as precision, recall, and F1-score were calculated. Finally, Table 3 presents a comparative analysis benchmarking the proposed system against recent LSTM-based fall detection methods, focusing on accuracy, precision, recall, and F1-score.

Table 3*. Comparative analysis benchmarking.*

Model	Accuracy	Precision	Recall	F1-Score
GRU	99.12%	1.00	0.98	0.99
LSTM	99.56%	1.00	0.99	1.00

Comparison Table 3 indicates that the LSTM-based processing model outperforms the GRU-based model by less than 1%. This suggests that while the LSTM model offers marginally better results, it is more cumbersome and computationally intensive than the GRU model. Consequently, it can be concluded that the benefits of using the LSTM model do not justify the additional computational costs, indicating limited computational efficiency. Moreover, this also demonstrates that the MediaPipe-GRU integration proposed in this paper differs from the MediaPipe-LSTM approach.

3. EXPERIMENTAL RESULTS

3.1 Physical hardware implementation

The fall detection system proposed in this paper utilizes the following hardware configuration for training and experimentation.

- Data acquisition utilized a Hikvision DS-2CV1021G0-IDW1 surveillance camera.
- Processing computer: The processing system is deployed on the Windows 11 platform, with hardware configuration including:
    - Processor: Intel(R) Core(TM) i5-1135G7 11th generation, speed 2.40 GHz.
    - RAM: 16 GB.
    - Graphics: Intel(R) Iris(R) Xe Graphics.

In addition, the design algorithms are operated on the following software platforms:

- Programming language Python 3.12.3.
- Development platform using PyCharm Community Edition.
- Framework and libraries: TensorFlow, NumPy, Pandas, Matplotlib, Scikit-learn, OpenCV.

3.2 Experimental scenarios and results

Our fall detection system was rigorously evaluated across two distinct age groups: adults and children. Real-time experiments demonstrated the system's effectiveness in continuously monitoring subjects and issuing immediate alerts upon detection of a fall. Results from these real-time experiments consistently exceeded 90% accuracy, coupled with a demonstrably favorable processing speed.

Finally, the experimental results are depicted in Figures 7 and 8.

Figure 7. No-Fall *detection*

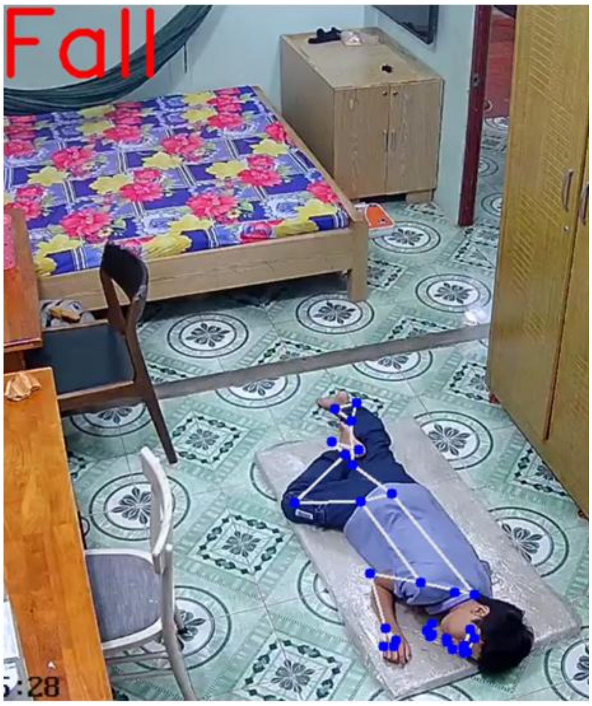

Figure 8. *Fall detection*

4. CONCLUSION

In conclusion, the fall detection system introduced in this paper marks a significant leap forward in leveraging computer vision to enhance safety for vulnerable populations, particularly the elderly. By employing advanced feature point analysis, our system achieves high accuracy and reliability in real-time fall detection, enabling timely interventions that can mitigate injury risks. This innovative approach underscores the efficacy of robust feature extraction and preprocessing techniques, effectively distinguishing falls from routine activities while minimizing false positives through a carefully designed algorithmic framework.

Our findings validate the system's potential as a cornerstone for intelligent monitoring in healthcare and smart home environments. The successful deployment of this technology lays a strong foundation for future enhancements, such as integrating adaptive machine learning models to further refine detection precision and robustness over time. As global aging populations grow, the demand for reliable, scalable safety solutions intensifies, positioning this system as a critical tool for fostering independence and well-being.

Looking ahead, we envision expanding the system's capabilities through integration with complementary smart technologies, creating holistic safety ecosystems. This research sets a precedent for computer vision-driven innovations that blend technical excellence with societal impact, paving the way for advancements in automated healthcare monitoring. We are confident that this work will inspire further exploration and development, ultimately improving the quality of life for those at risk of falls.

5. ACKNOWLEDGMENT

We acknowledge VNUHCM-University of Information Technology, and the Thu Dau Mot University (TDMU)'s Scientific for this study.

6. REFERENCES

[1] Wang X, Ellul J, Azzopardi G. "Elderly Fall Detection Systems: A Literature Survey." Front Robot AI. 2020 Jun 23;7:71. doi: 10.3389/frobt.2020.00071. PMID: 33501238; PMCID: PMC7805655.

[2] Namoos AM, Thomson N, Bradley S, Aboutanos M. "The Impact of Demographics and Comorbidities on Fall Incidence and Prevalence in Older Adults." Res Sq [Preprint]. 2024 Jul 19:rs.3.rs-4762014. doi: 10.21203/rs.3.rs-4762014/v1. PMID: 39070613; PMCID: PMC11276020.

[3] Deng, R., Li, B., Qin, M. et al. "The characteristics and risk factors of fatal falls among adults aged 60 and above in Southwest China." Sci Rep 14, 7020 (2024). https://doi.org/10.1038/s41598-024-54265-9.

[4] Ye P, Er Y, Wang H, Fang L, Li B, Ivers R, Keay L, Duan L, Tian M. "Burden of falls among people aged 60 years and older in mainland China, 1990-2019: findings from the Global Burden of Disease Study 2019." Lancet Public Health. 2021 Dec;6(12):e907-e918. doi: 10.1016/S2468-2667(21)00231-0. PMID: 34838197; PMCID: PMC8646839.

[5] Montero-Odasso M, van der Velde N, Martin FC, Petrovic M, Tan MP, Ryg J, Aguilar-Navarro S, Alexander NB, Becker C, Blain H, Bourke R, Cameron ID, Camicioli R, Clemson L, Close J, Delbaere K, Duan L, Duque G, Dyer SM, Freiberger E, Ganz DA, Gómez F, Hausdorff JM, Hogan DB, Hunter SMW, Jauregui JR, Kamkar N, Kenny RA, Lamb SE, Latham NK, Lipsitz LA, Liu-Ambrose T, Logan P, Lord SR, Mallet L, Marsh D, Milisen K, Moctezuma-Gallegos R, Morris ME, Nieuwboer A, Perracini MR, Pieruccini-Faria F, Pighills A, Said C,

Sejdic E, Sherrington C, Skelton DA, Dsouza S, Speechley M, Stark S, Todd C, Troen BR, van der Cammen T, Verghese J, Vlaeyen E, Watt JA, Masud T; task Force on Global Guidelines for Falls in Older Adults. “World guidelines for falls prevention and management for older adults: a global initiative.” Age Ageing. 2022 Sep 2;51(9):afac205. doi: 10.1093/ageing/afac205. Erratum in: Age Ageing. 2023 Sep 1;52(9):afad188. doi: 10.1093/ageing/afad188. Erratum in: Age Ageing. 2023 Oct 2;52(10):afad199. doi: 10.1093/ageing/afad199. PMID: 36178003; PMCID: PMC9523684.

[6] Nguyen TV, Tran HM, Trinh HBT, Vu VH, Bang VA. “Prevalence of frailty according to the Hospital Frailty Risk Score and related factors in older patients with acute coronary syndromes in Vietnam.” Australas J Ageing. 2024 Jun;43(2):288-296. doi: 10.1111/ajag.13307. Epub 2024 Apr 4. PMID: 38576179.

[7] Huyen BT, Van Anh PT, Duong LD, The THN, Guo P, Van Thuc P, Khue LN, Krakauer EL, Harding R. “Quality of life among advanced cancer patients in Vietnam: a multicenter cross-sectional study. Support Care Cancer.” 2021 Aug;29(8):4791-4798. doi: 10.1007/s00520-021-06012-3. Epub 2021 Feb 3. PMID: 33537888; PMCID: PMC8236448.

[8] Ministry of Health, "Preventing falls in the elderly" July 2, 2020. Electronic version: https://moh.gov.vn/web/phong-chong-tai-nan-thuong-tich/tin-noi-bat/-/asset_publisher/iinMRn208ZoI/content/phong-chong-te-nga-o-nguoi-cao-tuoi. [Accessed April 11th, 2025].

[9] Sadik Kamel Gharghan, Huda Ali Hashim, “A comprehensive review of elderly fall detection using wireless communication and artificial intelligence techniques,” Measurement, Volume 226, 2024, 114186, ISSN 0263-2241, https://doi.org/10.1016/j.measurement.2024.114186.

[10] Subramaniam S, Faisal AI, Deen MJ. “Wearable Sensor Systems for Fall Risk Assessment: A Review.” Front Digit Health. 2022 Jul 14;4:921506. doi: 10.3389/fdgth.2022.921506. PMID: 35911615; PMCID: PMC9329588.

[11] Chen M, Wang H, Yu L, Yeung EHK, Luo J, Tsui KL, Zhao Y. “A Systematic Review of Wearable Sensor-Based Technologies for Fall Risk Assessment in Older Adults.” Sensors (Basel). 2022 Sep 7;22(18):6752. doi: 10.3390/s22186752. PMID: 36146103; PMCID: PMC9504041.

[12] Orejel Bustos AS, Tramontano M, Morone G, Ciancarelli I, Panza G, Minnetti A, Picelli A, Smania N, Iosa M, Vannozzi G. “Ambient assisted living systems for falls monitoring at home.” Expert Rev Med Devices. 2023 Jul-Dec;20(10):821-828. doi: 10.1080/17434440.2023.2245320. Epub 2023 Aug 23. PMID: 37610096

[13] Chouhan, K., Kumar, A., Chakraverti, A. K., & Cholla, R. R.. “Human fall detection analysis with image recognition using convolutional neural network approach.” In Proceedings of Trends in Electronics and Health Informatics. pp. 95–106, 2022, Springer.

[14] Sykes, E. R. (2024). Next-generation fall detection: harnessing human pose estimation and transformer technology. Health Systems, 14(2), 85–103. https://doi.org/10.1080/20476965.2024.2395574.

[15] C. Ge, I. Y. -H. Gu and J. Yang, “Co-Saliency-Enhanced Deep Recurrent Convolutional Networks for Human Fall Detection in E-Healthcare,” 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 2018, pp. 1572-1575, doi: 10.1109/EMBC.2018.8512586.

[16] Sengar SS, Kumar A, Singh O. “Efficient Human Pose Estimation: Leveraging Advanced Techniques with MediaPipe.” arXiv:2406.15649.

[17] Google for Developers, "MediaPipe: Enhancing Virtual Humans to be more realistic," July 10, 2023. Electronic version: https://developers.googleblog.com/en/mediapipe-enhancing-virtual-humans-to-be-more-realistic/. [Accessed April 11th, 2025].

[18] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, Yoshua Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," Presented in NIPS 2014 Deep Learning and Representation Learning Workshop, https://doi.org/10.48550/arXiv.1412.3555.

[19] Google AI for Developers, "Pose landmark detection guide," Jan 13, 2025. Electronic version: https://ai.google.dev/edge/mediapipe/solutions/vision/pose_landmarker. [Accessed April 11th, 2025].