A New Method for Real-Time Fall Detection Based on MediaPipe Pose Estimation and LSTM

Puwadol Sirikongtham, Apichaya Nimkoompai
Faculty of Information Technology, Thai-Nichi Institute of Technology, Thailand

***Abstract*—Falls are a significant health problem among older adults, leading to serious injuries and adversely affecting both quality of life and public health burdens. Although various fall detection systems have been developed using technologies such as wearable sensors and image processing (computer vision), limitations remain in dimensions of convenience, accuracy, and real-time responsiveness. To overcome these limitations, this research aimed to present a real-time fall detection system that integrates MediaPipe pose estimation technology with a Long Short-Term Memory (LSTM) neural network. The proposed method functioned through two main components. MediaPipe pose estimation technology was applied to detect and track keypoints on the human body from real-time video input; meanwhile, a trained LSTM model was utilized to analyze the sequence of movements of the detected keypoints for classifying and differentiating between fall behaviors and normal activities. The system was trained, and its performance was evaluated using the standard UR Fall Detection Dataset. From experimental results, the proposed system achieved high efficiency in fall detection, with an accuracy of 95.2% on the test dataset. The integrated system had its capability to detect all actual fall events (with a recall of 100%). Its false positive rate was low. Compared to other research, the proposed method provided higher accuracy. These results indicated that the proposed system has the potential for practical application as an effective tool for real-time fall alerts, enabling timely assistance for those injured from falls.**

Keywords*—*Fall detection; older adults; real-time; MediaPipe; LSTM; computer vision; pose estimation

I. Introduction

Falls among older adults have long been recognized as a critical global health problem, leading to severe injuries, long-term disabilities, and even fatalities. According to the World Health Organization (WHO), approximately 28 to 35% of individuals aged 65 and above experience at least one fall annually, with 20 to 30% of these cases resulting in moderate to severe injuries [1]. Beyond the medical consequences, falls impose a substantial economic and social burden on families and healthcare systems worldwide [2]. As populations continue to age, the demand for effective and efficient fall detection solutions has become increasingly urgent, both to improve the quality of life for older adults and to alleviate the strain on healthcare infrastructures. Over the past decade, researchers have developed a variety of fall detection systems utilizing different technologies, such as wearable sensors and vision-based approaches. Wearable devices, while capable of capturing physiological and motion data with reasonable accuracy, are often uncomfortable for continuous use and may be resisted by older adults [3]. Conversely, vision-based systems, particularly those employing image processing and deep learning, can provide non-intrusive monitoring but frequently encounter challenges related to real-time responsiveness, computational efficiency, and accuracy under diverse environmental conditions [4]. In addition, many vision-based approaches rely on spatial features alone, without adequately modeling the temporal dynamics of human motion, which are essential for distinguishing fall incidents from similar activities such as sitting or bending [11]. To address these limitations, this study proposes a novel real-time fall detection method that integrates MediaPipe pose estimation with a Long Short-Term Memory (LSTM) neural network. MediaPipe enables lightweight and efficient detection of human body landmarks across multiple platforms [5] [6] [7], while LSTM is well-suited for learning temporal sequences of movement patterns [8] [9]. By combining these technologies, the proposed system aims to achieve high accuracy in fall detection while maintaining practical feasibility for real-time applications. Experimental evaluation was conducted using the UR Fall Detection Dataset, a widely adopted benchmark in the field [12], to ensure the reliability and comparability of results. The contributions of this study are threefold: 1) the development of a real-time fall detection framework that effectively integrates MediaPipe and LSTM for temporal motion analysis, 2) empirical validation on a standard dataset, demonstrating superior performance with an accuracy of 95.2% and a recall of 100%.

II. Related Work

A. MediaPipe Pose Estimation

Pose estimation technology was applied as a fundamental component of this research, particularly the MediaPipe Pose Estimation developed by Google Research. This was an end-to-end framework designed for real-time human pose detection and tracking. This technology utilized an optimized BlazePose architecture, enabling the detection and continuous tracking of up to 33 key body landmarks with high efficiency, processing at a rate of up to 30 frames per second on standard hardware [5].

The MediaPipe Pose Estimation system consisted of two main components. First, the BlazePose Detector was utilized to identify the initial location of the body within an image. Second, the BlazePose Landmark Model provided detailed localization of landmarks on the body. The study by Bazarevsky [6] showed that MediaPipe achieved high accuracy in both controlled and real-world environments with an accuracy rate of up to 87.8% on the MS Coco dataset. Furthermore, the system was capable of running on common mobile devices without requiring high-end GPUs.

Key advantages of MediaPipe Pose Technology include:

1) Real-time capability through specialized architecture optimization, enabling processing speeds fast enough for immediate application.

2) High accuracy, performing well even in conditions with partial obstructions or incomplete viewpoints.

3) Multi-platform support, capable of running on both mobile devices and general-purpose computers.

4) 3D coordinate calculation, in addition to 2D image pose estimation, it can also estimate positions in a 3D world.

In the context of fall detection, MediaPipe Pose Technology provides several crucial pieces of information, namely:

- The positions of various body joints
- Angles between key joints
- The speed and movement patterns of joints

The operation of MediaPipe Pose Technology begins with receiving an input image through the following steps:

1) Person Detection

2) Landmark Prediction

3) Post-processing

4) Output Generation by exporting results in 2D image coordinates and 3D world coordinates.

Numerous studies, such as that of Chiu [7], have proved that MediaPipe Pose achieved sufficient accuracy for applications in medical motion analysis and health-related monitoring tasks, while maintaining a processing speed that is suitable for real-time systems.

B. LSTM is a Type of Recurrent Neural Network (RNN)

LSTM is designed to address the issue of long-term dependencies. This is suitable for analyzing sequential data such as video or continuous human motion [8]. In this study, LSTM techniques were applied to learn the sequence of keypoints obtained from MediaPipe in order to classify fall behaviors among older adults. The LSTM function relied on its architecture, comprising a cell state, input gate, output gate, and forget gate, which collectively enable the network to retain long-term information and mitigate the vanishing gradient problem [9]. LSTM was capable of effectively analyzing motion patterns and learn the distinctive characteristics of individual falls.

C. Integration of MediaPipe and LSTM

When the camera captures video footage, the system utilizes MediaPipe to detect and track the positions of keypoints on the human body. These keypoints are then arranged in chronological order and fed into the LSTM to analyze movement patterns and classify whether an event is a fall. The integration of MediaPipe and LSTM allows the system to operate in real-time, with high accuracy and rapid response, which is critical in emergency situations. This design also helps to reduce the likelihood of false alarms and enhances the reliability of the system in real-world applications.

D. Related Research on Fall Detection

Fall detection has been a consistently significant topic in healthcare research, particularly among older adults who are at high risk of serious health consequences from falls. Various technologies have been applied to develop fall detection systems, including wearable sensors, computer vision systems, and artificial intelligence techniques to enhance behavior classification performance.

1) Sensor detection systems: The study by Kepski and Kwolek [10] developed a system using an accelerometer and a depth sensor to distinguish fall behaviors from normal activities. Although the system proved to have high accuracy, its reliance on specialized hardware limits its practical implementations. Similarly, Kulurkar [3] proposed a fall detection system that utilizes wearable sensors integrated with IoT technology in smart homes, enabling real-time alerts. However, this technique had limitations in terms of the inconvenience of wearing it for older adults.

2) Image and video processing system: Recently, vision-based systems have gained popularity, especially those incorporating deep learning techniques. Although the use of Convolutional Neural Networks (CNN) in the work by Saleh and Jeannès [11] has shown effectiveness in detecting falls from video due to CNN's capability in extracting spatial features from still images or individual video frames, a major limitation of this approach is the lack of temporal analysis. CNN treats video as a series of still images without accounting for the sequence or the transitional dynamics of pose before and after a potential fall. This temporal context is essential for distinguishing falls from other behaviors, such as sitting down or bending to pick something up.

The absence of temporal components may lead to confusion between similar movements that have different outcomes, such as, a rapid collapse caused by a fall versus an intentional quick sit-down. Without a model capable of learning from event sequences, the system could easily produce false positives or false negatives. Therefore, it is essential to utilize a model that can capture continuous movement patterns, such as LSTM, to enhance the accuracy of the analysis and ensure better contextual alignment with the actual event.

III. METHODOLOGY

A. System Architecture

The architecture of the proposed method is an integrated system of MediaPipe with an LSTM-based classifier to monitor and evaluate the movements of older adults in real-time. For its functionality, webcam video capture was utilized to continuously record motion data. Data were then processed using computer vision technology to extract pose data. The LSTM model has been trained on a dataset involving both general movement behaviors and fall incidents. These data were analyzed through the LSTM model to accurately detect fall events. The integration of these components is a key element of the proposed solution, ensuring effective hazard prevention and promoting the safety of older adults in their daily lives. The workflow diagram of the proposed system is shown in Fig. 1.

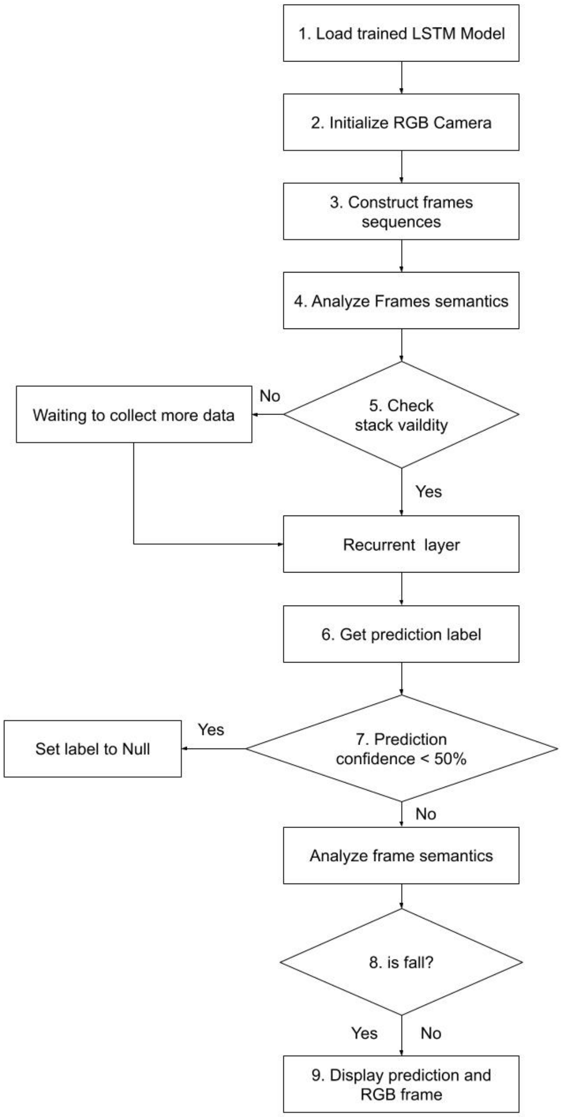

Fig. 1. Workflow diagram of the proposed real-time fall detection method using MediaPipe Pose Estimation and LSTM.

Real-Time Fall Detection System Workflow

1) Load the LSTM model along with its trained weights, which are capable of classifying movement behaviors (between normal activities and fall).

2) Initialize RGB camera by activating the webcam to capture real-time video input.

3) Generate frame sequence by extracting a sequence of consecutive video frames to gather sufficient motion data for analysis.

4) Analyze frames Using MediaPipe to extract pose landmarks of individuals in each frame.

5) Verify sequence accuracy to check whether the frame sequence is complete and sufficient.

- If incomplete, continue accumulating data.
- If complete, forward the sequence to the LSTM model for behavior classification.

6) Receive classification output from LSTM by obtaining a confidence score indicating whether the detected behavior corresponds to a "fall" or "non-fall" event.

7) Evaluate prediction confidence:

- If the confidence score is below 50%, label the prediction as null.
- If the confidence score exceeds 50%, proceed with further analysis.

8) Analyze the classification result by assessing whether the movement corresponds to normal behavior or a fall incident.

9) Display prediction results by presenting the classification result ("fall" or "non-fall") along with the relevant video frames for monitoring or alert.

B. Dataset Details

For developing this system, the dataset was derived from the UR Fall Detection Dataset [12], which consists of a total of 70 video clips. These are divided into 30 videos showing fall incidents and 40 videos capturing normal movement behaviors. The videos were recorded from various camera angles and perspectives. Participants with different heights and physical characteristics were included to reflect a broad range of real-world motion patterns.

During the data preparation process, MediaPipe was applied to extract keypoints from each video frame, as shown in Fig. 2. These keypoints were then stored in a clearly structured folder. The dataset covers sequences of various movement behaviors, including walking, sitting, and different types of falls, making it well-suited for training the LSTM model to effectively learn and detect fall events. Detailed characteristics of the dataset are presented in Table I.

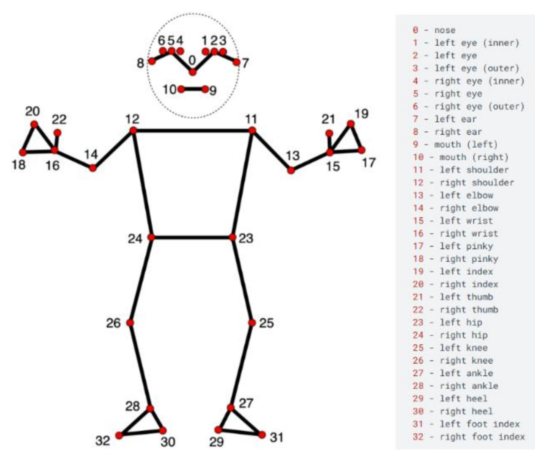

Fig. 2. Example of 33 human body landmarks identified using MediaPipe.

C. Implementation

In the initial phase of system implementation, real-time human pose detection was conducted using the MediaPipe Holistic model. Images were continuously captured via an RGB camera and processes to extract landmark data of various body parts, such as the pose, face, and hands. These landmark data are considered key attributes for fall detection.

Long Short-Term Memory (LSTM) neural network is the system's core, which was designed to effectively learn time-sequenced data. The model incorporates multiple LSTM layers followed by Dense layers, which classify whether the user belongs to a "fall" or "non-fall" condition. This model is trained using the Adam optimizer and utilizes Categorical Cross-Entropy as the loss function.

A key feature of the LSTM layers is their ability to retain and learn behavioral patterns or motion sequences over a continuous period of time. This capability makes them well-suited for distinguishing between normal behaviors (such as walking or sitting) and abnormal events such as falling. The internal mechanism of the LSTM relies on gates that manage the flow of information within the memory units during each time step. These two components, including the landmark extraction and the LSTM-based classification, are driven by their different loss functions.

TABLE I COMPOSITION OF DATASET DETAILS

Posture Type	Falling	Normal
Number of Videos	30	40
Frames per Video	30	30
Landmarks per Frame	33	33
Total Keypoint Values Per Frame	132	132

Total Keypoint Values per Frame = Number of Landmarks × Number of Coordinate Values per Landmark Therefore, 33 Landmarks × 4 values per Landmark = 132 values per frame.

In this research, the MediaPipe Pose model utilized 3D coordinate values (x, y, z) + a confidence score, totaling four values per keypoint. These values were used as inputs for the LSTM model to learn pose characteristics. The confidence score assisted in filtering out uncertainly detected landmarks. A total of 33 keypoints (including 11 keypoints located in the head region) were used in the analysis. These points are significant to identify abnormal movement patterns, such as tilting or sudden head-dropping motions, which are often the initial signs of falls.

In this study, the following equations show the mechanism of the Long Short-Term Memory (LSTM) cells applied to analyze movement sequences extracted via MediaPipe:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t]) + b_i \tag{1}$$

Input Gate [Eq. (1)] controls the amount of new information from the current input ($\mathrm{x_t}$) and the previous hidden state ($\mathrm{h_{t-1}}$) that will be recorded in the cell state. A value close to 1 indicates a high allowance of information flow, whereas a value close to 0 blocks the information.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t]\right) + b_f \tag{2}$$

Forget Gate [Eq. (2)] determines whether to retain or discard the old information in the cell state (C_{t-1}). This eliminates irrelevant data, such as normal posture prior to a fall. A value close to 1 indicates the retained information, whereas a value close to 0 indicates the forgotten information.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t]) + b_o \tag{3}$$

Output Gate [Eq. (3)] controls which part of the cell state is passed on as the hidden state (h_t). This determines which aspects of the movement are important for fall detection.

$$C_t = f_t * C_{t-1} + i_t * \tanh(W_C \cdot [h_{t-1}, x_t] + b_C \tag{4}$$

Cell State Update [Eq. (4)] controls that irrelevant information is discarded via the Forget Gate (f_t), and new information is added via the Input Gate(i_t). The tanh function generates candidate values (ranging from –1 to 1) for updating the state, enabling the recording of movement patterns that may lead to a fall.

$$h_t = o_t * \tanh(C_t) \tag{5}$$

Hidden State [Eq. (5)] is the output of the LSTM cell that is passed to the next layer. It retains crucial posture information for prediction in the subsequent frame.

$$y = softmax(W \cdot h + b) \tag{6}$$

Classification Output [Eq. (6)] converts the final Hidden State into the probability of each class (fall / non-fall).

where,

x_t = Current input (keypoint values from MediaPipe at time t)

h_t = Hidden State (information passed between cells)

C_t = Cell State (long-term memory)

σ = Sigmoid (Squashing Function)

tanh = Hyperbolic Tangent

W = Weight Matrix learned to capture motion patterns from MediaPipe data.

b = Bias used to balance the activation of each gate. The fall detection system was trained using sequence data over 150 epochs. The Early Stopping technique was implemented to monitor the categorical accuracy. Training was automatically terminated once the accuracy stabilized, in order to prevent overfitting, which may arise from excessive learning on the training data.

After completion of the training process, the model was evaluated using a test dataset comprising 21 videos (12 normal activity videos and 9 fall incidents), accounting for 30% of the total dataset. The evaluation involved the generation of a confusion matrix and the calculation of an accuracy score to assess the model's ability to classify behaviors in videos it had not encountered during training.

For extracting pose features, the system utilized the MediaPipe Pose model, which can detect 33 keypoints on the human body. Among these, 11 keypoints located around the head were included due to their critical role in analyzing abnormal movements such as sudden tilting or dropping of the head which are often early signs of falls.

In real-world application, the system captures real-time video input through an RGB camera, processes it using MediaPipe to detect keypoints, and then feeds the sequence data into the pre-trained LSTM model to predict behavior. The results are displayed immediately on screen, both as text and audio alerts, to help caregivers promptly detect and respond to fall incidents.

IV. Results

A. Experimental Results

The experimental results of the fall detection system are presented in Table II, which shows the performance metrics of the developed model. The results indicate that the model was capable of accurately classifying behaviors, with only one misclassification out of a total of 21 samples. The test dataset consisted of 9 videos of actual falls (positive class) and 12 videos of normal activities (negative class). The system achieved an accuracy of 0.952, with most false positives caused by rapid sit-stand movements or sudden bending to pick up objects.

To further illustrate the system's performance Fig. 3 and Fig. 4 provide visual examples of the MediaPipe Pose Estimation and LSTM model's outputs. Fig. 3 depicts the successful detection and classification of a normal posture, where MediaPipe accurately tracks 33 keypoints, including critical head and body joints, and the LSTM model assigns a high-confidence "Normal Posture" label. In contrast, Fig. 4 shows the detection of a falling posture, capturing the rapid descent and tilt of keypoints.

TABLE II TRAINING AND EVALUATION INDICATORS OF THE MODEL

Metric	Value	Comments
Epochs	150	Total training epochs on the UR Fall Detection dataset
Accuracy	0.952	Accuracy on the test dataset (including both fall and non-fall classes)
False Positives	1	1 False Alarm (Normal Activity)
True Negatives	11	Correctly identified normal activities 11 times
True Positives	9	All falls were successfully detected
False Negatives	0	No missed fall detections
Precision	0.9	Precision of the alert
Recall	1	Fall detection rate
F1-Score	0.947	Balance between Precision and Recall

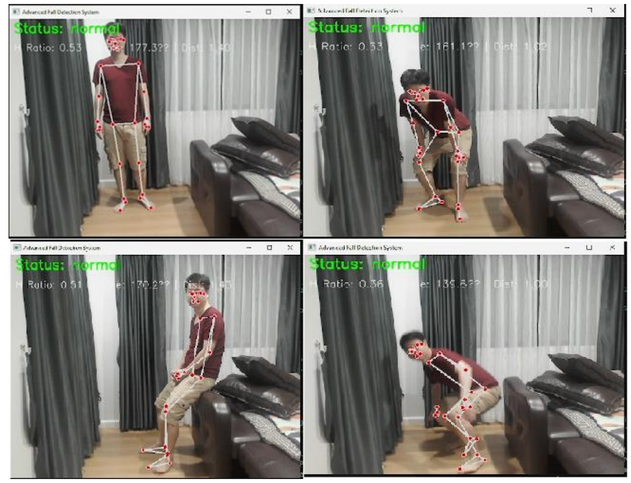

Fig. 3. Testing the MediaPipe Pose Estimation and LSTM model for normal posture.

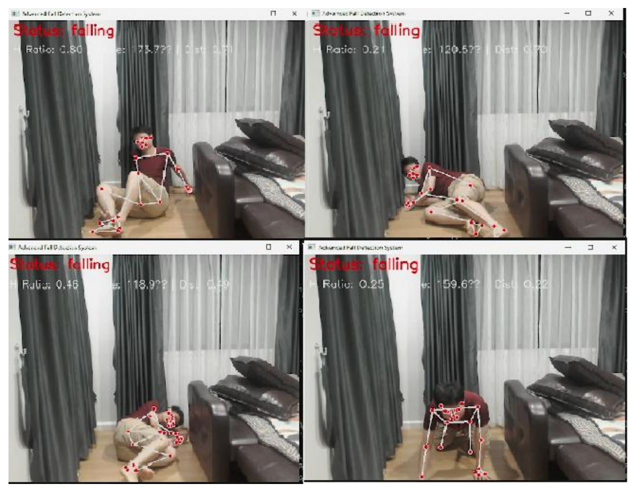

Fig. 4. Testing the MediaPipe Pose Estimation and LSTM Model for falling posture.

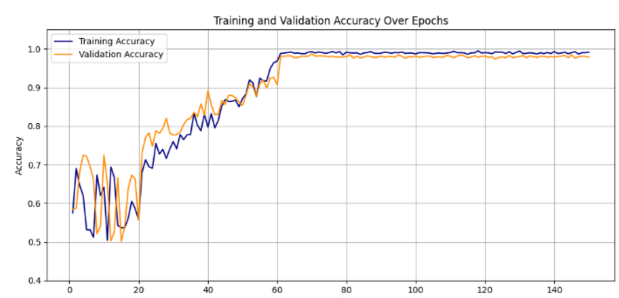

Fig. 5. Training and validation accuracy performance over epochs.

As shown in Fig. 5, the graph illustrates the model's accuracy throughout the training period, revealing a stable and steadily increasing learning trend, as the number of epochs increases.

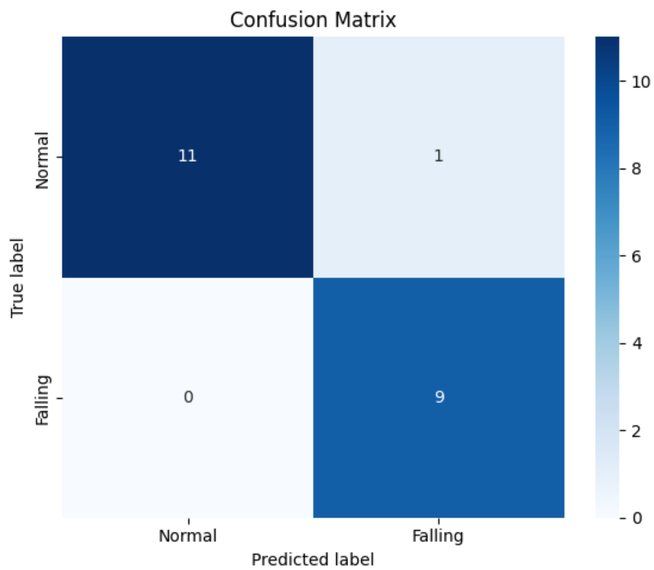

Fig. 6. Confusion matrix of the system.

Fig. 6 presents the confusion matrix, which displays the number of correctly classified instances (True Positives and True Negatives) and misclassified instances (False Positives and False Negatives). This allows for a detailed analysis of the system's error distribution. The fall detection system exhibited high performance when tested on 21 video clips, achieving a 100% recall, correctly identifying all fall incidents. The overall accuracy was 95.2%. There was only one false alarm from normal activities.

V. Discussion

A. *Comparative Analysis of Pose-Based Fall Detection Models*

In the development of a real-time fall detection system utilizing an LSTM neural network, comparative analysis with existing research is essential to prove the effectiveness of the proposed method. In this study, a comparative analysis was conducted with two related works. First, the research entitled "Real-Time Human Fall Detection using a Lightweight Pose Estimation Technique" [13], proposed a real-time fall detection system employing a lightweight pose estimation technique. The study utilized the Movenet model for extracting body keypoints. This system is capable of operating on resource-constrained devices without relying on cloud-based processing, thereby addressing privacy problems. The model achieved an accuracy of 91.67% on the URFD dataset [12].

Second, the research entitled "PreFallKD: Pre-Impact Fall Detection via CNN-ViT Knowledge Distillation" [14], proposed a pre-impact fall detection system by applying a knowledge distillation technique between Vision Transformer (ViT) and Convolutional Neural Network (CNN) models. This approach created a high-performance model that is also compatible with resource-constrained devices. The system achieved an accuracy of 92.66%.

TABLE III COMPARATIVE RESULTS OF DIFFERENT MODELS

Model	Latency (ms/frame)	Accuracy
MediaPipe Pose + LSTM	33 ms	0.952
MoveNet (Lightweight CNN) [13]	15 ms	0.9167
CNN + Vision Transformer (ViT) [14]	50 ms	0.9266

From Table III, the comparison of fall detection accuracy across different real-time models was presented. The proposed MediaPipe + LSTM model in this study achieved the highest accuracy at 95.2%, outperforming both previous works. The model's latency was measured at 33 ms/frame, which qualifies as real-time performance (≤50 ms), though slightly slower than the Movenet-based method.

The results of the research entitled "Real-Time Human Fall Detection using a Lightweight Pose Estimation Technique" [13] were compared, which used the Movenet model and achieved the lowest latency of only 15 ms/frame due to its single-frame structure, and an accuracy of 91.67%. It was found that despite the advantages in speed and processing on low-resource devices, the foresaid system may have limitations in learning complex event sequences or behaviors due to the lack of a temporal sequential model like an LSTM.

Meanwhile, the results of the research entitled "PreFallKD: Pre-Impact Fall Detection via CNN-ViT Knowledge Distillation" [14] achieved an accuracy of 92.66%, which was also high. However, it had the highest latency at 50 ms/frame. Its strengths lie in its pre-impact detection capability and the effective use of knowledge distillation between Vision Transformer (ViT) and CNN. Nonetheless, this model had a more complex structure and demanded more resources for training compared to the MediaPipe + LSTM system.

In summary, the MediaPipe + LSTM model offers significant advantages, including the highest accuracy among all three models (95.2%) and a latency of 33 ms/frame, which falls within the real-time threshold. It effectively learned temporal patterns of continuous poses using LSTM and efficiently extracted keypoints with MediaPipe. Furthermore, it does not require advanced hardware or cloud processing.

From these comparative experimental results, it can be proven that the proposed method was highly efficient and had significant potential for pose detection and fall risk assessment. This can be applied in nursing homes or hospitals to enhance safety and quality of life.

VI. Conclusion

In this research, an innovative real-time fall detection method was proposed utilizing human pose estimation via MediaPipe Pose in conjunction with a Long Short-Term Memory (LSTM) neural network model. The integrated technique was capable of effectively learning sequential pose data and accurately distinguished fall events. Experimental results indicated that the proposed model was functional with high accuracy in classifying both normal activities and fall incidents, exhibiting strong accuracy and a low misclassification rate.

Enhancing the diversity of training data such as varying camera angles, lighting conditions, and pose characteristics of different samples should be emphasized for future directions to improve the system's adaptability to a wider range of real-world environments. In addition, further optimization of the model should be promoted for faster performance and a more lightweight architecture. This is for the implementation with resource-constrained devices, such as standard CCTV cameras or IoT devices.

Lastly, plans should be established to improve the accuracy and responsiveness of the alert system, including connecting the system to applications for caregivers or medical personnel. This aims to enhance the effectiveness of fall prevention and care for older adults or individuals at risk. A key future goal is to evaluate the long-term effectiveness of the system through collaborations with public health agencies, in order to assess its impact on user safety and quality of life.

REFERENCES

[1] World Health Organization, "Falls," 26 April 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/falls. [Accessed: May . 30, 2025].

[2] World Health Organization, *WHO Global Report on Falls Prevention in Older Age*. Geneva, Switzerland: WHO, 2007, p. 4. [Online]. Available: https://extranet.who.int/agefriendlyworld/wp-content/uploads/2014/06/WHo-Global-report-on-falls-prevention-in-older-age.pdf). [Accessed: May . 30, 2025].

[3] P. Kulurkar, C. K. Dixit, V. C. Bharathi, A. Monikavishnuvarthini, A. Dhakne, and P. Preethi, "AI-Based Elderly Fall Prediction System Using Wearable Sensors: A Smart Home-Care Technology with IoT," Measurement: Sensors, vol. 25, p. 100614, Feb. 2023, doi: 10.1016/j.measen.2022.100614.

[4] P. Sonawane, M. Patil, S. Shinde, and R. Patil, "Live Object Recognition Using YOLO," International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 13, no. 4, pp. 1308–1311, Apr. 2025. [Online]. Available: https://www.researchgate.net/publication/380207553_Live_Object_Recognition_Using_YOLO.

[5] J. W. Kim, J. Y. Choi, E. J. Ha, and J. H. Choi, "Human Pose Estimation Using MediaPipe Pose and Optimization Method Based on a Humanoid Model," Applied Sciences, vol. 13, no. 4, Art. no. 2700, 2023. [Online]. Available: https://www.mdpi.com/2076-3417/13/4/2700.

[6] V. Bazarevsky, I. Grishchenko, K. Raveendran, F. Zhang, and M. Grundmann, "BlazePose: On-device Real-time Body Pose Tracking," arXiv preprint arXiv:2006.10204, 2020. [Online]. Available: https://arxiv.org/abs/2006.10204.

[7] Y. H. Chiu, L. C. Lin, R. C. Yang, and Y. P. Chang, "An Automatic Approach of Video-Based Landmark Detection and Movement Analysis for Assessing Symptoms of Bradykinesia in Parkinson's Disease," Journal of Medical and Biological Engineering, vol. 44, 2024. [Online]. Available: https://link.springer.com/article/10.1007/s40846-024-00876-6.

[8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[9] F. Ordóñez and D. Roggen, "Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition," Sensors, vol. 16, no. 1, p. 115, 2016, doi: 10.3390/s16010115.

[10] M. Kępski and B. Kwolek, "Human fall detection on embedded platform using depth maps and wireless accelerometer," Computer Methods and Programs in Biomedicine, vol. 117, no. 3, pp. 489–501, 2014, doi: 10.1016/j.cmpb.2014.09.005.

[11] B. Saleh and R. Jeannès, "Vision-Based Fall Detection Using Convolutional Neural Networks," Computers in Biology and Medicine, vol. 111, 2019, doi: 10.1016/j.compbiomed.2019.103355.

[12] M. Kępski, UR Fall Detection Dataset, Interdisciplinary Centre for Computational Modelling, University of Rzeszow. [Online]. Available: https://fenix.univ.rzeszow.pl/~mkepski/ds/uf.html.

[13] E. Alam, A. Sufian, P. Dutta, and M. Leo, "Real-Time Human Fall Detection using a Lightweight Pose Estimation Technique," arXiv preprint arXiv:2401.01587, Jan. 2024. [Online]. Available: https://arxiv.org/abs/2401.01587.

[14] T.-H. Chi, K.-C. Liu, C.-Y. Hsieh, Y. Tsao, and C.-T. Chan, "PreFallKD: Pre-Impact Fall Detection via CNN-ViT Knowledge Distillation," arXiv preprint arXiv:2303.03634, Mar. 2023. [Online]. Available: https://arxiv.org/abs/2303.03634.