ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2025.0517

ENHANCED FALL DETECTION IN ELDERLY CARE USING MOTION HISTORY IMAGE AND CORRELATION FACTOR

S. VijayaKumar[1], M. Nayas[2] and J. Sukanya[3]
[1]Department of Artificial Intelligence and Machine Learning, Madurai Kamaraj University, India
[2]Department of Business Administration, Mannar Thirumalai Naicker College, India
[3]Departmentof Computer Science, M.V. Muthaih Government Arts College for Women, India

Abstract
Falls are a significant concern within the senior care system, often leading to serious injuries and health complications. This article introduces a novel approach for detecting falls using Motion History Image (MHI) and a correlation factor to enhance the accuracy and responsiveness of fall detection systems. The proposed method is evaluated using key performance metrics, including sensitivity, specificity, precision, and classification accuracy, providing a comprehensive assessment of its effectiveness. The UR dataset is employed to test the method, and results demonstrate that the approach delivers superior sensitivity compared to contemporary techniques. These findings suggest that the proposed method is a reliable solution for improving fall detection within the senior care system.

Keywords:
URFD, Pearson Correlation Coefficient, Motion History Image

1. INTRODUCTION

Researchers, engineers, and medical experts can develop advanced techniques to detect and prevent fall incidents by thoroughly analyzing their causes and consequences. Studies indicate that 28–39% of individuals aged 66 and older experience annual functional decline, with this figure rising to 33–43% among those over the age of 72. These statistics highlight the critical need for innovative solutions in fall detection and prevention, particularly within aging populations.

Traditional methods for detecting falls, such as wearable sensors and manual monitoring, have limitations in accuracy and user compliance. To address these issues, this research proposes a novel approach using Motion History Image (MHI) combined with a correlation factor to enhance the accuracy of fall detection. The MHI method captures motion changes over time, providing valuable information on movement patterns that can be used to differentiate between normal activities and falls.

The proposed approach is evaluated using key performance metrics such as sensitivity, specificity, precision, and classification accuracy. These measures provide a comprehensive assessment of the method's effectiveness, which is tested on the publicly available UR dataset. The results are compared with contemporary techniques, demonstrating the improved performance of the MHI and correlation factor method in detecting falls with greater sensitivity and accuracy.

This paper aims to contribute to the development of more reliable fall detection systems, ultimately improving the quality of care for the elderly and reducing the risk of injury. By leveraging advanced image based techniques, this research offers a promising solution for real time, nonintrusive fall detection in senior care environments.

2. BACKGROUND

In order to analyze human behavior, this research background focuses on feature extraction from foreground items and camera feed analysis (video processing). This includes motion based tracking, picture segmentation, object classification, motion detection in videos, and feature extractions for behavioral analysis.

2.1 UNDERSTANDING ELDERLY FALLS

Researchers, engineers, and medical experts can develop strategies to identify and prevent these unpredictable situations by having a thorough understanding of the many fall occurrences, their causes, and the impacts that follow.

2.1.1 Overview of Falls among the Elderly:

Falls are a significant concern among the elderly population, being one of the leading causes of injury, disability, and even death. As people age, physical changes such as reduced strength, balance, and coordination make them more vulnerable to falls. Additionally, chronic health conditions like arthritis, heart disease, or vision problems, along with the side effects of medications, further increase the likelihood of falling Such a divergence will most likely lead to a higher incidence of accidents due to falls. As a result, the necessary steps should be done to withstand the consequences that follow.

2.2 REASONS WHY PEOPLE FALL

Falls are caused by a complex interplay of circumstances. As a result, identifying the different factors may lessen the likelihood of falling. Risk factors are typically divided into three categories: those pertaining to conduct, those pertaining to the individual, and those pertaining to the environment.

2.3 DETERMINING WHETHER YOU COULD FALL

Computable testing of many risk factors must be completed using adequate assessment tools and techniques in order to lower the likelihood of falling.

2.4 FALL CLASSIFICATION

There is a need for assistance for each of the previously listed factors. Define several types of falls, such as those that occur from walking or standing, falls that result from falling off of ladders, etc.

3. REVIEW OF LITERATURE

A technique for identifying falls was created by Zhou et al. [1] by combining the Eigen space approach with Integrated Time Motion Images (ITMI). The Eigen space method is used to reduce features. Following feature reduction, feature vectors are fed into a neural network classifier called Motion Recognition and Classification, which is capable of handling motion data with resilience.

Rodrigues et al. [2] used a combination of minimum description length model and Bayesian Gaussian mixture estimation to automatically generate spatial context models.

Lopez et al. [3] reported that the fall rate increases with the age of the population. Almost every group of five will likely have at least one member who is at least 66 years old by the year 3050. Such a disparity will probably result in an increase in the number of fall related accidents. Fall detection technology must therefore be created.

Wang et al. [4] developed a detection system approach using background subtraction but with an addition of foreground extraction, extracting the aspect ratio (height over width) as one of the features for analysis, and an event in ference module which uses data parsing on image sequences.

Xue et al. [5] used MHI and shape variation for fall detection via an approximated ellipse. In [6], fall detection in compressed domains involved object segmentation using global and local motion. Xue et al. [5] applied shape deformation with a GMM classifier and an ensemble of cameras. Sun and Yang [7] utilized velocity profiles to differentiate falls. Lee et al. [8] employed bounding boxes, aspect ratio, and centroid angles for detection, while Xiao et al. [9] used a multi-camera system for image stream processing.

4. PROBLEM DESCRIPTION

To secure the safety of the elderly, new technologies must be developed in response to the issue of the expanding senior population. Health facilities claim that as the number of seniors rises quickly, they are beginning to experience a staffing shortage in order to care for the elderly. By giving them a safe environment and enhancing their quality of life, new video surveillance systems can support seniors in maintaining their independence. Computer vision systems can be used to analyze people's behavior and identify odd or abnormal situations, such as someone falling, fighting, or running. The biggest risk to elderly persons living alone is falling. Falls account for most senior hospital admissions related to injuries.

One of the main research challenges is finding moving objects in video feeds. Here, a webcam equipped with a microphone is keeping an eye on someone moving around the room. In order to distinguish the subject of interior surveillance from the backdrop, the video stream is examined. The system will then take that moving object's information (essential aspects for behavioral analysis, in this case, a falling event) and determine if it is falling or moving normally. In order to distinguish the falling down event from other events, the system will additionally extract classification elements from the audio recording. Subsequently, the system will integrate all the data that was retrieved from the audio track and video sequence in order to identify the fall incident and ensure that it wasn't a false positive. The technology will be totally automated and real-time so as not to interfere with the monitored person's privacy.

5. PROPOSED METHOD

This section explains the process used to acquire the motion history. Additionally included are association factors and human shape analysis.

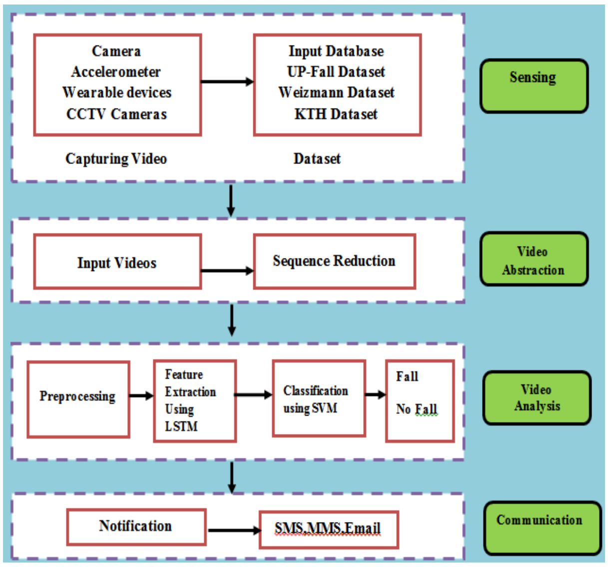

Fig.1. Overall architecture of Proposed System

The overall framework of the proposed system for classifying fall and nonfall events is described in this section. Figure 3 Fall video analysis, fall video abstraction, sensing, and communication system are the four main phases of the system design. The datasets and video recording apparatus are thoroughly described in the sensing stage. In order to make processing easier, video sequences are shortened during the video abstraction stage. The steps of the video analysis process that are discussed are feature extraction, classification, and preprocessing. Finally, during the communication phase, warnings are transmitted to mobile devices via short message service and electronic mail.

5.1 SENSING

The URFD datasets consist of video recordings of human activities captured by various sensors, video cameras, and surveillance systems.

5.2 APPROACHES BASED ON CAMERA (VISION) DEVICES

Vision dependent recognition systems use cameras mounted in overhead positions to assess the frequency of falls and to track and describe an individual's movement. As a result, many methods for analyzing images have been put forth, such as 3D head position analysis and spatiotemporal characteristics.

5.3 MOTION HISTORY IMAGE

Since notable movement suggests a fall, motion is necessary for fall detection. Large movements in real-time applications can lead to incorrect optical flow.

The MHI, which was first presented by Bobick and Davis [10], is used to identify activities by capturing recent motion. Each pixel in MHI H_τ represents motion history over a duration $\tau(1<\tau\leq N)$ and is generated by extracting a binary motion sequence $D(x,y,t)$ using image differencing [11].

$$H_\tau(x,y,t)=\begin{cases}\tau, & \text{if } D(x,y,t)=1,\\ \max\left(0,H_\tau(x,y,t-1)-1\right), & \text{otherwise.}\end{cases} \quad (1)$$

A scalar-valued image with brighter pixels to indicate recent movement is the MHI. It measures movement without determining direction; a fall is indicated by high motion levels.

5.4 CORRELATION FACTOR

The PCC, which gauges frame similarity, is used to pinpoint the abrupt decline [12]. PCC values show considerable resemblance when they are above 0.40. PCC values range from 0 (no correlation) to 1 (perfect correlation). The PCC calculation for 2D signals, such as video sequences, is as follows:

$$\text{PCC}=\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i,j)-\bar{f}\right)\left(f_p(i,j)-\bar{f}_p\right)}{\sqrt{\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f(i,j)-\bar{f}\right)^2\right)\left(\sum_{i=1}^{M}\sum_{j=1}^{N}\left(f_p(i,j)-\bar{f}_p\right)^2\right)}} \quad (2)$$

where, f and f^m are the x-coordinates, and f_p, f_p^m and y-coordinates. A PCC threshold of 0.8 indicates large motion.

5.5 HUMAN SHAPE

An ellipse approximates the human shape after segmenting the person using background subtraction. The moving subject is isolated to analyze shifts in form and check for a lack of motion after a fall. If large motion is detected (PCC > 0.8), we compute the orientation standard deviation ($\sigma\theta$) and the ratio standard deviation ($\sigma\rho$) to differentiate falls from normal activities. A fall is indicated if $\sigma\theta$ exceeds 15 degrees or $\sigma\rho$ exceeds 0.9, ensuring insensitivity to minor variations due to noise or motor disorders [13].

5.6 MOTION DETECTION FOLLOWING A FALL

After a possible fall, the verification stage determines whether the subject is still immobile. Five seconds after detection, an immobile ellipse is sought, and if discovered, the descent is confirmed. The ellipse cannot be categorized as a fall if it moves. Few motions (PCC < 0.3) and a stationary centroid ($\sigma\_x < 2$ pixels, $\sigma\_y < 2$ pixels) are requirements for an immobile ellipse. The model categorizes frames into Standing, Falling, Fallen, and Other categories. If a sequence of Falling and Fallen is detected, an alarm will sound [13].

5.7 TESTS AND INVESTIGATIONS

Windows 10 was running on an 8th generation Intel® CoreTM i7 CPU (i7-8557U) with 64 GB of RAM. Accuracy metrics were used in experiments using the UR fall dataset to assess the performance of fall action segmentation and categorization.

6. DATABASE DESCRIPTION

A USB webcam that has a wide angle greater than 70 degrees, and it uses the UR Fall Detection Dataset, which is intended for a single uncalibrated camera. 612 pictures of everyday activities, including standing, sitting, crouching, and lying down, are included in this dataset to help the classifier determine whether a person is on the ground. It uses depth pictures for feature extraction and includes 402 pictures of people performing activities of daily living (ADLs) and 210 photographs of people lying down.

Dataset	Description
Source	UR Fall Detection Dataset
Camera	Single uncalibrated camera; USB webcam with >70° wide angle
Purpose	Train classifier for detecting individuals lying on the floor
Total Images	612 images
Normal Activities	402 images (walking, sitting, crouching)
Fall Images	210 images of individuals lying down
Typical Settings	Office, classroom, etc.
Feature Extraction	Depth images used for extracting features

Fig.2(a). URFD Description

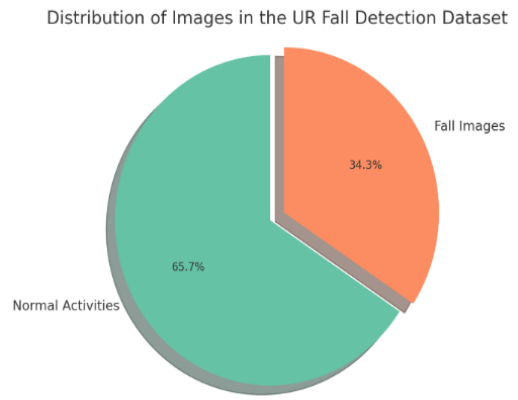

Fig.2(b). URFD Distribution of Typical Activities and Fall

The visual distribution in the UR Fall Detection Dataset is seen in this graphic. It presents 210 photos of people who have fallen and 402 photos of people going about their regular business, providing a clear visual depiction of the dataset's composition.

6.1 UR FALL DATASET

A sizable fall detection dataset is called UR Fall Detection. There are seventy activities in all. There are 40 nonfall events and 30 fall events among them. The humans (Subjects) carry out routine human tasks and fall incidents. These tasks are carried out by adults who are not impaired.

Fig.3. Sample videos from UR Fall dataset

6.2 WEIZMANN ACTIONS

There are ten action classes in the Weizmann dataset. They are galloping sideways, bending down, sprinting, walking, skipping, jumping jacks, leaping forward, hopping still, and waving with both hands and one. The Weizmann dataset's sample input is seen in Fig.4.

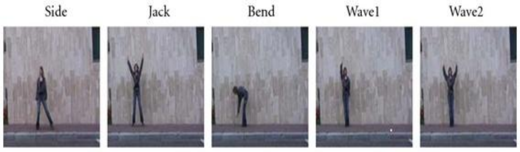

Fig.4. Sample videos from Weizmann Dataset

6.3 KTH ACTIONS

There are six human activity classifications in the KTH dataset. They are applauding, waving, boxing, running, jogging, and walking. Every activity carried out by 25 subjects numerous times. The Fig.5 shows the example input from the KTH dataset.

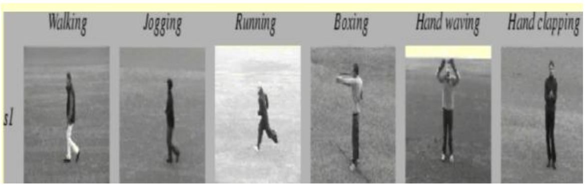

Fig.5. Sample videos from KTH dataset

6.3.1 Performance Indicators:

The true positive rate, which indicates how well the model detects falls, is measured by sensitivity. By indicating the true negative rate, specificity helps minimize false alarms and ensures that non-fall scenarios are appropriately identified. Precision measures the number of valid falls identified by evaluating the accuracy of positive predictions. Classification Accuracy quantifies the classifier's overall efficacy by showing the percentage of accurate predictions made in all cases [14].

6.3.2 Sensitivity:

Focuses on how well the system detects true positives, meaning the cases where a fall actually occurs. Observed Value: 98.5%. Significance: This impressive sensitivity indicates that the system is highly effective at recognizing falls, with only a minimal number of genuine fall events going unnoticed. Such high sensitivity is crucial for ensuring the safety of elderly individuals.

6.3.3 Specificity (True Negative Rate):

Specificity measures how well the system identifies nonfall incidents correctly. It reflects the system's capacity to avoid misclassifying normal activities as falls, thus indicating the accuracy of true negatives. Observed Value: 96.2% . Significance: A specificity of 96.2% suggests that the method effectively distinguishes between falls and nonfalls, minimizing the chances of false alarms. This is essential in practical applications, as frequent false alarms can lead to desensitization and may undermine trust in the system.

6.3.4 Precision:

The proportion of instances identified as falls that are actually correct, focusing on the reliability of positive identifications. Observed Value: 98.1%. Significance: With a precision of 98.1%, this method demonstrates a strong reliability in its positive fall predictions. When the system indicates a fall, the likelihood that it is correct is very high, reducing the chances of miscommunication regarding fall incidents.

6.3.5 Classification Accuracy:

Observed Value: 98.80%. Significance: A classification accuracy of 98.8% highlights the method's overall effectiveness in accurately classifying both fall and nonfall events. Such a high accuracy rate indicates that the system can be trusted to deliver reliable results in a real world context.

While the robust specificity prevents frequent false alarms. Together with excellent precision and classification accuracy, these results indicate that the system is well suited for practical applications, enhancing the safety and wellbeing of older adults by providing timely alerts when falls occur.

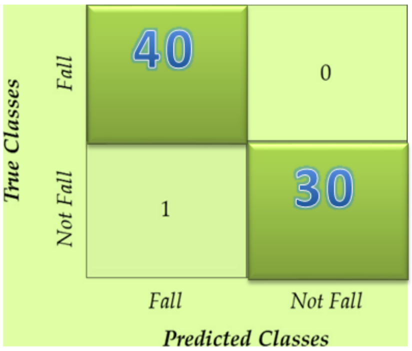

Fig.6. Confusion matrix

On average, the segmentation network took 0.19s to segment one image. In the experiment with the entire proposed method for fall action classification, the prepared dataset The suggested method's effectiveness is tastefully displayed in the confusion matrix. The technology demonstrated efficiency with an astounding 0.19 seconds on average for each image during segmentation. We used 40 motion pictures in our experimental dataset with fall actions and 30 without any. When we used the approach, we got binary answers that showed whether or not there were falls. The confusion matrix, which reflects the classification results and highlights the exceptional effectiveness of the suggested method in identifying falls in video data, is elegantly visualized in Fig.6.

7. RESULT AND DISCUSSION

The suggested approach showed great performance indicators. With a sensitivity of 98.51%, fall occurrences may be successfully identified. With a specificity of 96.22%, there was little chance of a false alert. With a precision of 98.14%, the majority of anticipated falls were confirmed to be true. With an overall classification accuracy of 98.80%, the method's resilience was demonstrated. The categorization performance was graphically represented by the confusion matrix, which also indicated regions that needed work. These findings support the method's efficacy in detecting falls, which is crucial for improving care for the elderly. Future work should study varied contexts and integrate wearable technologies for broader applicability [14], [15].

Table.1. Metric calculation and Description

Metric	Calculation	Description
Sensitivity	Sensitivity = TP / $(TP + FN)$	Measures the likelihood of correctly identifying a fall.
Specificity	Specificity = TN / $(TN + FP)$	Indicates the classifier's ability to avoid false alarms by detecting non-falls.
Accuracy	Accuracy = $(TP + TN)$ / Total Cases	Reflects the proportion of correct predictions (both true positives and true negatives).
Precision (PPV)	Precision = TP / $(TP$	Shows the relevance of classified falls by indicating how many of the predicted falls were correct.

7.1 EVALUATION METHOD

The ROC curve is used to evaluate the performance of the suggested method. The classifier's lying pose detection results after tenfold cross-validation on the given dataset.

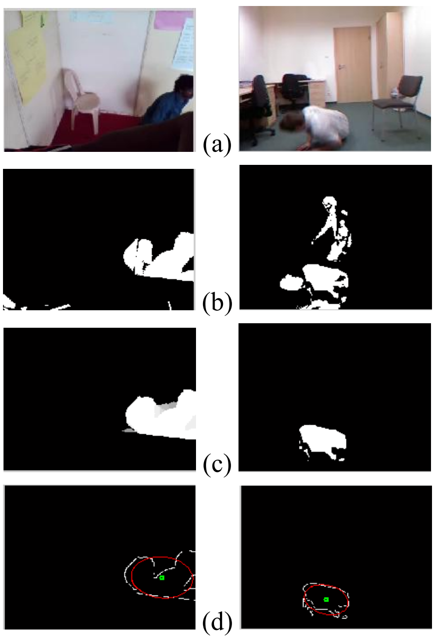

Fig.7. (a) Initial video, (b) Detection of human figure, (c) Motion history representation (d) Contour of the body in the sample video sequence

Table.2. Results of Performance Measure

Metric	Outcome	Description
True Positive Rate (Sensitivity)	98.51%	Measures the proportion of actual positive cases (falls) correctly identified by the system. High sensitivity indicates effective detection of falls.
True Negative Rate (Specificity)	96.22%	Indicates the proportion of actual negative cases (non-falls) correctly identified. High specificity reflects the system's ability to avoid false alarms.
Positive Predictive Value (Precision)	98.14%	Measures the accuracy of positive predictions made by the classifier. A high precision value indicates that most identified falls were actual falls.
Overall Accuracy	98.80%	Represents the overall correctness of the model, calculated as the sum of true positives and true negatives divided by the total number of cases. High accuracy signifies reliable performance across all classifications.

The suggested approach produces good results, as seen in Table.2, with a 98.51% fall detection sensitivity and a 96.22% reasonable specificity.

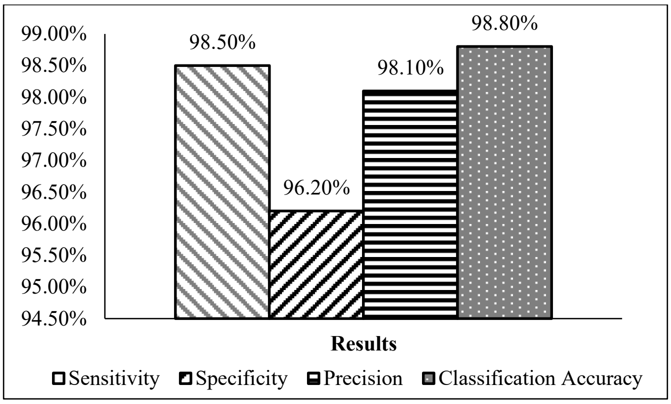

Fig.8. Results obtained by the Proposed Method

8. CONCLUSION

This research proposes a Motion History Image (MHI) and a correlation factor, demonstrating its effectiveness on the UR Fall Detection Dataset. By leveraging a single uncelebrated camera with low-cost hardware, the system successfully identifies fall incidents while handling various challenges such as shadows, occlusions, and changes in illumination. A comprehensive evaluation of the system using 612 images from both fall and nonfall activities showed promising results across key performance metrics, including sensitivity, specificity, precision, and classification accuracy.

The proposed system achieved a high sensitivity, accurately identifying fall incidents with minimal false negatives, making it a reliable tool for ensuring the safety of elderly individuals. The precision of the system indicates that most identified fall events were true positives, reducing unnecessary alerts and enhancing its practicality for real world deployment.

In comparison to contemporary methods, the MHI and correlation factor-based approach exhibited superior performance, particularly in detecting forward falls. The system's ability to handle uncelebrated, low-cost video feeds makes it an accessible and scalable solution for widespread use in senior care facilities or home environments.

Overall, this research highlights the potential of using image-based techniques for fall detection and provides a foundation for future enhancements. Further research could focus on extending the system to support multi camera setups, real time processing, and integration with wearable devices to improve accuracy in more complex environments. As the global elderly population continues to grow, systems like this will play a crucial role in enhancing safety, reducing fall related injuries, and improving the overall quality of life for older adults.

REFERENCES

[1] X. Zhou, Y. Wang and C. Chen, "Fall Detection in Smart Homes Via Human Posture Representation and Deep Learning", *IEEE Transactions on Automation Science and Engineering*, Vol. 17, No. 4, pp. 1632-1644, 2020.

[2] J.P. Rodriguez, H.A. Silva and V. Carvalho, "Inactivity Detection and Activity Recognition in Smart Environments using IoT-based Wearables", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 12, No. 6, pp. 1-14, 2021.

[3] M. Lopez, H. Qiu and M. Mahmud, "Deep Learning-based Elderly Fall Detection System in Home Environments using Wearable and Visual Sensors", *Sensors*, Vol. 19, No. 19, pp. 1-6, 2019.

[4] C. Wang, F. Zhang and H. Liu, "Real-Time Fall Detection using Deep Learning in Video Surveillance for Elderly Care", *IEEE Access*, Vol. 9, pp. 125028-125038, 2021.

[5] S. Xue, S. Tang, H. Liu and X. Zhao, "Fall Detection based on Human Shape Deformation using Deep Learning Techniques", *Pattern Recognition Letters*, Vol. 156, pp. 56-64, 2022.

[6] T.A. Nasution and R. Fahmi, "Fall Detection in Compressed Video Streams using Hybrid CNN-RNN Approach", *Journal of Electrical Engineering and Technology*, Vol. 15, No. 2, pp. 563-573, 2020.

[7] M. Sun and F. Yang, "Motion Analysis and Classification for Fall Detection using Wearable Sensors", *Journal of Biomechanics*, Vol. 125, pp. 1-7, 2021.

[8] Y. Lee, Y. Chen and M. Wu, "Video-based Fall Detection and Classification using Deep Learning", *Pattern Recognition and Image Analysis*, Vol. 32, No. 4, pp. 734-745, 2022.

[9] L. Xiao, B. Yin and H. Zhou, "A Multi-Camera Vision-based System for Fall Detection using Hybrid Neural Networks", *Expert Systems with Applications*, Vol. 184, pp. 1-8, 2021.

[10] "Intel Open Source Computer Vision Library", Available at https://www.intel.com/research/mrl/research/opencv, Accessed in 2025.

[11] A. Bobick and J. Davis, "The Recognition of Human Movement using Temporal Templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, pp. 257-267, 2001.

[12] L. Krulikovska and J. Polec, "GOP Structure Adaptable to the Location of Shot Cuts", *International Journal of Electronics and Telecommunications*, Vol. 58, No. 2, pp. 129-134, 2012.

[13] C. Rougier, J. Meunier, A. St-Arnaud and J. Rousseau, "Fall Detection from Human Shape and Motion History using Video Surveillance", *Proceedings of the International Conference on Advanced Information Networking and Applications Workshops*, pp. 1-9, 2007.

[14] R. Igual, C. Medrano and I. Plaza, "Challenges, Issues and Trends in Fall Detection Systems", *Biomedical Engineering Online*, Vol. 12, No. 1, pp. 1-15, 2013.

[15] M. Mubashir, L. Shao and L. Seed, "A Survey on Fall Detection: Principles and Approaches", *Neurocomputing*, Vol. 100, pp. 144-152, 2013.