JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

AI-BASED FALL DETECTION MONITORING SYSTEM

REAL-TIME PROTECTION FOR ELDERLY AND PATIENT SAFETY

Dr. B. Rajesh Kumar, Nithin Ganesh. C Assistant Professor, Student
Department of Software Systems and AIML
Sri Krishna Arts and Science College, Coimbatore, India

Abstract **:** Falls are a major safety concern for elderly people and patients, especially in hospitals, care facilities, and home environments where continuous human observation may not always be possible. This paper proposes an AI-based fall detection monitoring system that uses computer vision and deep-learning techniques to identify possible fall events from video input. The system captures video frames, performs preprocessing, extracts visual and motion-related features, and classifies activities as normal activity or fall. When a fall is detected, the system can generate an immediate alert and display the detection result for caregivers or medical personnel. The proposed approach is designed to support continuous monitoring, reduce response time, and improve safety while reducing dependence on manual observation. The system can be extended with real-time camera monitoring, event recording, notification services, and integration with healthcare environments. Performance can be evaluated using accuracy, precision, recall, F1-score, false-alarm rate, and detection time.

IndexTerms **- Artificial Intelligence, Fall Detection, Elderly Safety, Patient Monitoring, Computer Vision, Deep Learning, Video Surveillance, Real-Time Alert.**

I. INTRODUCTION

Falls are one of the important safety risks faced by elderly people and patients. A fall can result in injuries, loss of consciousness, fractures, or other medical emergencies, and delayed assistance can increase the seriousness of the incident. Continuous observation by caregivers is useful but may be difficult in hospitals, care homes, and residential environments because personnel cannot monitor every location continuously.

Artificial Intelligence (AI), computer vision, and deep-learning techniques provide an opportunity to automate fall monitoring from camera footage. An intelligent system can analyze consecutive video frames, identify changes in body posture and movement, and determine whether an observed activity is likely to be a fall. The objective is not to replace caregivers, but to provide an additional monitoring and alert mechanism that can support faster human intervention.

The proposed system processes video from a camera or uploaded video source. Frames are extracted and preprocessed before a trained AI model analyzes spatial appearance and movement patterns. The final output identifies normal activity or a possible fall and can trigger an alert when a fall is detected.

A. Background

Traditional patient and elderly monitoring generally depends on direct observation, CCTV monitoring, wearable devices, or manual reporting. These methods can be useful, but manual observation can be time-consuming and a camera feed may be difficult for one person to monitor continuously. A computer-vision-based system can assist by automatically screening video and highlighting possible fall events.

Fall detection is challenging because normal activities such as sitting, lying down, bending, walking, and quickly changing posture may resemble parts of a fall. The system therefore needs to consider temporal movement and posture changes rather than relying only on a single image.

B. Problem Statement

How can an AI-based monitoring system automatically identify possible falls of elderly people and patients from video footage and generate a timely alert while minimizing false alarms during normal activities?

C. Objectives

1. Develop an AI-based fall detection framework for elderly and patient safety.
2. Process video frames using computer-vision techniques and extract relevant visual and motion features.
3. Classify observed activities into normal activity and fall events.
4. Generate an alert when a possible fall is detected.
5. Evaluate the system under different lighting, camera-angle, distance, and background conditions.
6. Measure detection performance using standard classification and response-time metrics.

D. Research Questions

RQ1: Can an AI-based video monitoring system detect possible falls from real-time or recorded video?

RQ2: Can deep-learning and computer-vision features distinguish fall events from normal activities?

RQ3: Can automated alerts support faster caregiver response after a detected fall?

RQ4: How do lighting, camera position, occlusion, and background conditions affect fall detection performance?

RQ5: How can false alarms be reduced while maintaining reliable fall detection?

II. RELATED WORK

Deep-learning approaches such as Convolutional Neural Networks (CNNs) have been widely used for visual recognition tasks. CNN-based feature extraction can identify body appearance, posture, and scene information from individual video frames. Temporal models such as Long Short-Term Memory (LSTM) networks can analyze sequences of extracted features and are suitable for activities in which the order and duration of movements are important.

Object-detection approaches such as YOLO can identify people and objects in video frames. EfficientNet is another deep-learning architecture that can be used for image feature extraction with an efficient model design. OpenCV provides tools for video loading, frame extraction, image preprocessing, and camera access.

Fall detection remains challenging because different environments produce different camera views, body orientations, illumination conditions, and levels of occlusion. A reliable system therefore requires representative training data and testing under varied conditions.

III. PROPOSED METHODOLOGY

The proposed system consists of video input, preprocessing, representative frame extraction, feature extraction, temporal analysis, classification, confidence-score generation, and alert output. The complete process is designed to analyze movement over time instead of making a decision from a single frame.

A. Video Input and Loading

Video can be obtained from a surveillance camera, webcam, or an uploaded video file. The system uses OpenCV to open the video stream and read frames sequentially. Basic validation is performed to ensure that the input can be processed.

B. Representative Frame Extraction

Instead of processing every available frame at the same computational cost, representative frames can be selected from the video sequence. The selected frames preserve important changes in body position and movement that may occur before, during, and after a fall.

C. Frame Preprocessing

Each selected frame is resized to the input dimensions expected by the AI model and normalized where required. Preprocessing helps maintain consistent input quality and can reduce the effect of differences in image size and pixel scale.

D. CNN Feature Extraction

A CNN is used to extract spatial features from individual frames. These features can represent information such as the person's appearance, body posture, position, and surrounding scene. A suitable pretrained architecture such as EfficientNet can be considered for feature extraction.

E. LSTM Temporal Analysis

The extracted frame features are arranged as a sequence and provided to an LSTM-based temporal analysis stage. The LSTM examines changes across successive frames and learns temporal patterns associated with movement. This helps distinguish a rapid change in posture associated with a possible fall from ordinary activities such as walking or sitting.

F. Classification and Confidence Score

The temporal representation is passed to a classification layer that predicts the activity class. The system can use two primary classes: Normal Activity and Fall. A confidence score is generated with the prediction so that the result can be displayed to the user and used as part of the alert logic.

G. Alert Generation

When the predicted class indicates a possible fall, the system generates an alert containing the detection status and time. The alert can be displayed on the monitoring interface and may be extended to a notification mechanism for caregivers or medical staff. Human verification is recommended for important medical or emergency decisions.

IV. DATASET CONSTRUCTION

A controlled dataset should contain video sequences representing both normal and fall activities. Normal examples may include walking, sitting, standing, lying down, bending, and other routine movements. Fall examples should represent different fall directions, speeds, body positions, and environmental settings where legally permissible data are available.

The dataset should include variations in lighting, camera angle, camera distance, background, clothing, and partial occlusion. Data should be divided into training, validation, and testing sets so that the final evaluation measures performance on previously unseen samples.

Table 1. Activity classes

Normal Activity — No fall or emergency movement is observed.

Fall — A person falling, collapsing, or rapidly moving from an upright position toward the floor is detected.

V. SYSTEM ARCHITECTURE

The system architecture consists of a camera or video source, video-processing module, AI-based feature extraction module, temporal analysis module, activity classifier, result display, and alert mechanism. The camera provides the input, the processing pipeline analyzes the sequence, and the output communicates the predicted activity to the responsible user.

System Flow

START → Video Input → Video Loading → Video Preprocessing → Representative Frame Extraction → Frame Preprocessing → CNN Feature Extraction → Feature Sequence Formation → LSTM Temporal Analysis → Fall / Normal Classification → Confidence Score → Display Result → Alert if Fall Detected → END

VI. EXPERIMENTAL DESIGN

The system can be evaluated through model comparison, real-time detection, environmental testing, alert evaluation, normal-activity detection, and fall-event detection. The experiments should measure both the ability to identify falls and the ability to avoid unnecessary alerts during normal activities.

Environmental testing should include different lighting conditions, camera angles, distances, backgrounds, and levels of partial occlusion. Misclassified samples should be reviewed to identify common causes such as poor illumination, rapid movement, overlapping people, low video quality, and unusual body postures.

A. Evaluation Metrics

Accuracy measures the proportion of correctly classified samples. Precision measures the proportion of predicted fall events that are actually fall events. Recall measures how many actual fall events are successfully detected. F1-score combines precision and recall. False-alarm rate measures unnecessary fall alerts, while detection time measures the time between the occurrence of a fall and generation of the alert.

VII. RESULTS AND DISCUSSION

The proposed system should be evaluated using the prepared test dataset and recorded performance metrics. Results can be presented using accuracy, precision, recall, F1-score, confusion matrix, false-alarm rate, and average detection time. The results should also be reported separately for normal activities and fall events so that missed falls and unnecessary alerts can be analyzed.

Performance may vary according to activity type and environmental conditions. In particular, low lighting, camera-angle changes, occlusion, complex backgrounds, rapid movement, and multiple people in the scene can affect detection. Therefore, evaluation should include these conditions rather than relying only on controlled examples.

VIII. ADVANTAGES AND APPLICATIONS

1. Supports continuous monitoring of elderly people and patients.
2. Provides automated identification of possible fall events.
3. Can reduce the need for continuous manual observation.
4. Can generate alerts to support faster human response.
5. Can be used with recorded video as well as real-time camera input.
6. Can support hospitals, elderly-care facilities, assisted-living environments, and home monitoring.

IX. ETHICAL CONSIDERATIONS

Video-based patient and elderly monitoring involves privacy and safety considerations. Video data should be collected and processed only with appropriate authorization and according to applicable institutional and legal requirements. Access to recorded footage should be restricted to authorized personnel.

AI predictions should not be treated as definitive medical diagnoses. False positives and false negatives are possible, so important incidents should be verified by an appropriate human responder. The system should be presented as a safety-support tool rather than a replacement for professional medical care.

X. LIMITATIONS

The performance of fall detection can vary with camera position, lighting, video quality, body orientation, occlusion, background complexity, and the number of people in the scene. Similar movements such as sitting quickly or lying down may sometimes resemble a fall. A larger and more diverse dataset and extensive real-world testing are therefore required before deployment in safety-critical environments.

XI. CONCLUSION

This paper presents an AI-based fall detection monitoring system designed to support elderly and patient safety. The proposed pipeline combines video preprocessing, representative frame extraction, CNN-based spatial feature extraction, LSTM-based temporal analysis, activity classification, confidence-score generation, and alert output. By analyzing movement patterns across video frames, the system can provide automated assistance in identifying possible fall events.

The system is intended to complement human monitoring by providing timely detection and notification. Future evaluation should focus on accuracy, recall, false-alarm reduction, detection time, robustness across environments, and responsible handling of video data.

XII. FUTURE ENHANCEMENT

Future work can extend the system toward real-time multi-camera monitoring, mobile and web notifications, larger and more diverse datasets, improved deep-learning architectures, edge-device deployment, event history, and integration with healthcare alert systems. Additional activity classes can also be introduced to distinguish falls from other emergency

or abnormal events.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Sri Krishna Arts and Science College, Coimbatore, for providing the academic support and facilities required for this work. The authors also thank the faculty members and project guides for their guidance and encouragement.

REFERENCES

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[2] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Proceedings of the International Conference on Machine Learning (ICML), 2019.

[3] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proceedings

of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[5] W. Liu et al., "SSD: Single Shot MultiBox Detector," European Conference on Computer Vision (ECCV), pp. 21–37, 2016.
[6] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.