The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 4, No 2, February 2026
PP 9–19

Vision-Based Fall and Mobility Risk Detection in Hospitalized Patients

Bejugam Vardhan Sai[1*] , Akula Ramya Sree[2], Ramakrishna Gandi[3]
[1,2]School Bejugam Vardhan Sai UG Scholar, Dept of CSE(AI&ML), Geethanjali College of Engineering and Technology, Hyderabad, Telangana, India
[3] Assistant Professor, Department of Computer Science & Engineering(AI&ML),Geethanjali College of Engineering and Technology, Hyderabad, Telangana ,India
Corresponding Author *: bejugamvardhansai@gmail.com

Abstract

Ensuring patient safety in healthcare environments requires reliable detection of fall incidents under diverse conditions such as occlusion, lighting variations, and dynamic human movement. This work presents a vision-based patient monitoring framework that combines human pose estimation with deep learning techniques to identify falls and abnormal activities. The system extracts skeletal key points from video data and utilizes a Bidirectional Gated Recurrent Unit (Bi-GRU) model to capture temporal motion patterns. By analyzing sequential pose features, the model distinguishes between normal and critical activities. Experimental results demonstrate an accuracy of 80%, precision of 83%, recall of 76%, and F1-score of 79%. The proposed approach offers a non-contact, scalable, and efficient solution for real-time patient monitoring. Future enhancements will focus on multi-modal data integration and system optimization for deployment.

Keywords:
human pose estimation, fall detection, patient monitoring, computer vision, deep learning, GRU, healthcare monitoring systems.

1. Introduction

Patient monitoring has become a critical requirement in modern healthcare systems, especially for elderly and high-risk individuals who are more prone to falls and abnormal movements. Falls in hospital environments can lead to serious injuries, increased hospitalization time, and higher medical costs, making early detection essential for improving patient safety.

Traditional monitoring methods rely on manual supervision or wearable sensors. However, these approaches have limitations such as delayed response, dependency on human observation, and discomfort to patients. In addition, wearable-based systems may fail due to improper usage, reducing reliability.

With advancements in computer vision and deep learning, vision-based monitoring systems have emerged as a promising solution. These systems analyse human movements using video data and enable continuous, non-intrusive monitoring without requiring wearable devices.

Research Objectives and Methodology

This study aims to develop a vision-based patient monitoring system using BlazePose and deep learning techniques. The objectives are:

- To detect fall events using human pose estimation.
- To extract 33 body landmarks for accurate posture representation.
- To apply a Bidirectional GRU model for temporal motion analysis.
- To improve real-time monitoring and detection accuracy.

2. Literature Survey

Early studies on fall detection relied on sensor-based systems such as accelerometers and pressure sensors, providing moderate accuracy but requiring wearable devices [1]. These systems often caused discomfort and depended on proper usage, limiting their effectiveness [2].

With technological advancements, alternative approaches such as Wi-Fi and radar-based monitoring were introduced, enabling contactless detection. However, these methods require specialized hardware and are sensitive to environmental conditions [3].

Computer vision-based approaches improved monitoring systems by using pose estimation techniques such as OpenPose to extract skeletal key points from video data [4]. These methods enhanced activity recognition but often lacked temporal analysis.

Deep learning models such as GRU and LSTM have been used to analyse sequential motion patterns, improving fall detection accuracy [5]. Hybrid approaches combining pose estimation with temporal modelling provide better performance in real-time monitoring systems [6]. However, challenges such as occlusion, lighting variations, and computational complexity still remain.

3. Methodology

3.1 Data Collection

A video-based dataset was used for patient monitoring, consisting of normal and abnormal activities, including fall events. The dataset contains multiple samples of human motion sequences labelled for activity classification.

3.2 Preprocessing

- **Video Frames:** Extracted and resized to fixed dimensions.
- **Pose Data:** Skeletal key points normalized for consistency.
- **Noise Reduction:** Background variations minimized.

3.3 Feature Extraction

- **Pose Estimation:** Extracts body key points (x, y coordinates) from each frame.
- **Feature Vector:** Each frame represented as:

  $X = [x_1, y_1, x_2, y_2, \ldots, x_n, y_n]$
- **Sequence Input:** Temporal sequence formed as:

  $Xt = \{X_1, X_2, \ldots, X_t\}$

3.4 Temporal Modeling (GRU)

- **GRU Model:** Processes sequential data to capture motion patterns.
- **Hidden State Update:**

  $h_t = GRU(X_t, h_{t-1})$
- **Output:**

  $y = f(h_t)$

3.5 Detection Mechanism

- **Decision Rule:**

  $y = 1 \rightarrow$ Fall detected

  $y = 0 \rightarrow$ Normal activity

- **Threshold:** Defined during validation for classification.

3.6 Evaluation

Split: 70% training, 20% validation, 10% testing.

Metrics:

- **Accuracy:**
  Accuracy = (TP + TN) / (TP + TN + FP + FN)
- **Precision:**
  Precision = TP / (TP + FP)
- **Recall:**
  Recall = TP / (TP + FN)
- **F1-Score:**
  F1 = 2 × (Precision × Recall) / (Precision + Recall)

models.

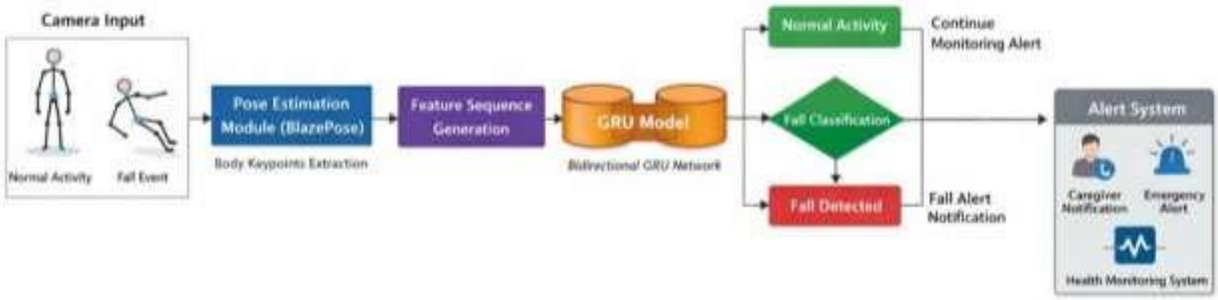

Fig.1 Bidirectional GRU Model for Vision Based Fall Detection

4. Experimental Setup and Implementation

4.1 Hardware Configuration

- **Processor:** Intel i5 / i7 or equivalent multi-core processor.
- **Memory:** 8 GB (minimum), 16 GB recommended.
- **GPU:** NVIDIA GPU (optional for faster training).
- **Storage:** SSD for efficient data processing.
- **OS:** Windows 10/11 or Ubuntu.

4.2 Software Environment

- **Language:** Python 3.8+.
- **Framework:** TensorFlow, Keras.
- **Libraries:** NumPy, Pandas, OpenCV, MediaPipe, Matplotlib, Seaborn.
- **Interface:** Gradio for real-time monitoring.
- **Development:** Kaggle Notebook / Local IDE.

4.3 Dataset Preparation

- **Data:** Video-based dataset containing normal and fall activities.
- **Pose Extraction**: BlazePose extracts 33 body landmarks (x, y, z) per frame.
- **Features:** 99 features per frame (33 landmarks × 3 coordinates).
- **Sequence Formation:** 30-frame sequences used for temporal modelling.
- **Split:** 70% training, 20% validation, 10% testing.

4.4 Training Process

- **Model:** Bidirectional GRU (Bi-GRU).
- **Architecture:** Two GRU layers (128, 64 units) with dropout and batch normalization.
- **Batch Size:** 32.
- **Epochs:** ~20 (early stopping applied).
- **Training:** Sequential pose data used for classification (Fall / Normal).
- **Optimization:** Learning rate scheduling and regularization applied.

4.5 Hyperparameter Tuning

- **GRU Units:** Tested (64–128).
- **Epochs:** Optimized using early stopping.
- **Learning Rate:** 0.001 (tested 0.0001–0.01).
- **Dropout:** 0.35 to reduce overfitting.

4.6 Baseline Implementation

- **Sensor-Based Systems:** Accelerometer-based monitoring.
- **Traditional Vision Methods:** Frame-based posture detection.
- **Comparison:** Evaluated against proposed pose + GRU system.

4.7 Evaluation Setup

- **Metrics:** Accuracy, Precision, Recall, F1-score.
- **Additional Tools:** Confusion matrix, ROC curve.
- **Visualization:** Graphs using Matplotlib and Seaborn.
- **Real-Time Testing:** Webcam-based evaluation using Gradio interface.

5. Result Analysis

the results analysis evaluates the performance of the proposed vision-based patient monitoring system using a test dataset consisting of normal and fall activity sequences. The Bidirectional GRU (Bi-GRU) model analyses temporal pose data extracted using BlazePose and classifies activities into fall and non-fall categories. The evaluation is based on standard classification metrics such as accuracy, precision, recall, and F1-score.

Test set results demonstrate the effectiveness of the proposed system:
- Accuracy: 0.80, indicating that 80% of the total predictions are correct.
- Precision: 0.83, showing that the model correctly identifies fall events with high reliability and low false positives.
- Recall: 0.76, indicating the model’s ability to detect most of the actual fall events.
- F1-Score: 0.79, providing a balanced measure of precision and recall.

The training performance is further analysed using loss convergence:
- Initial Loss: 0.62
- Final Loss: 0.12
- Loss Reduction Rate = (0.62 − 0.12) / 20 = 0.025

Table 1. Performance Metrics

Model / Method	Accuracy	Precision	Recall	F1-Score
Machine Learning (DT / KNN / SVM)	0.72	0.70	0.69	0.69
Pose-Based Deep Learning Model	0.78	0.76	0.75	0.75
Proposed Bi-GRU Model	0.80	0.83	0.76	0.79

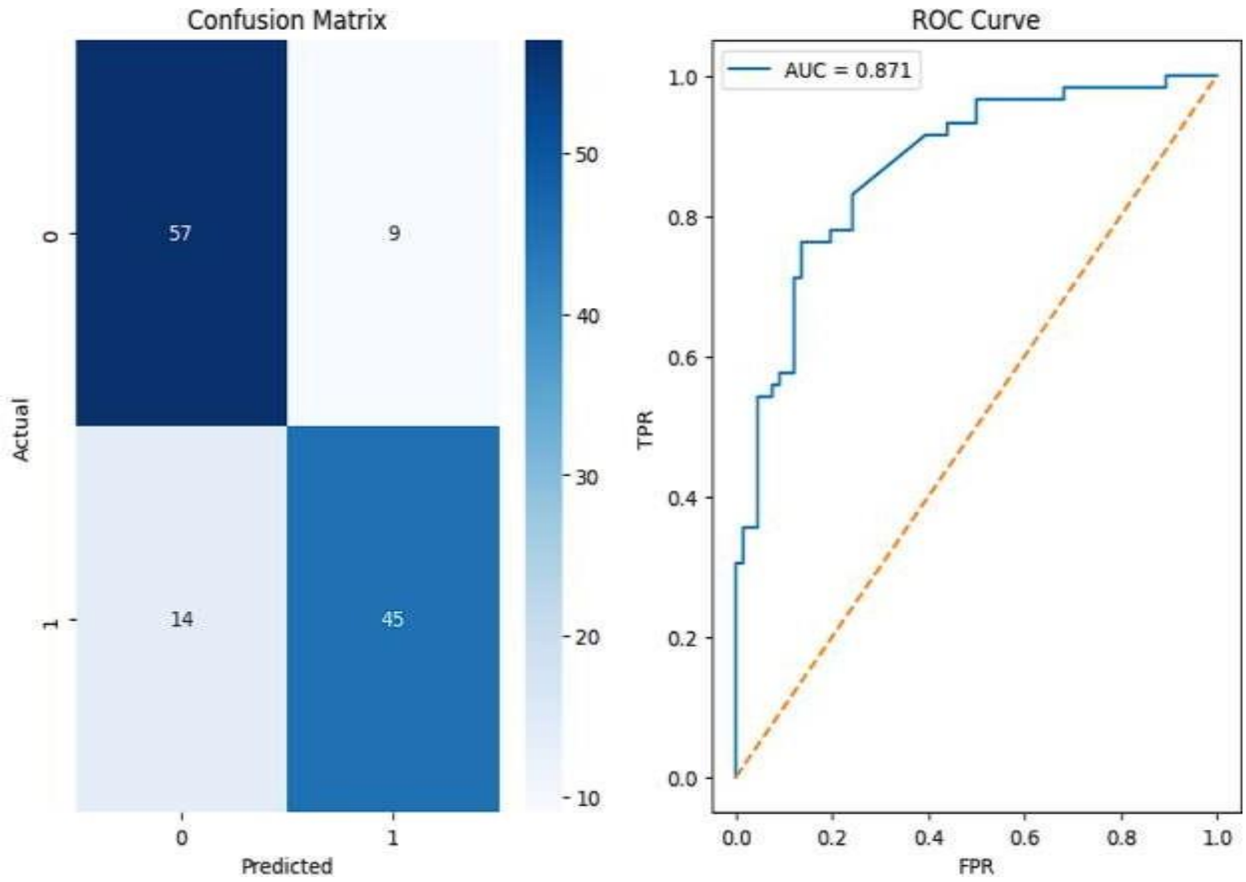

Fig2. Confusion Matrix of the Proposed GRU Model & ROC Curve and AUC Score for Fall Detection Performance.

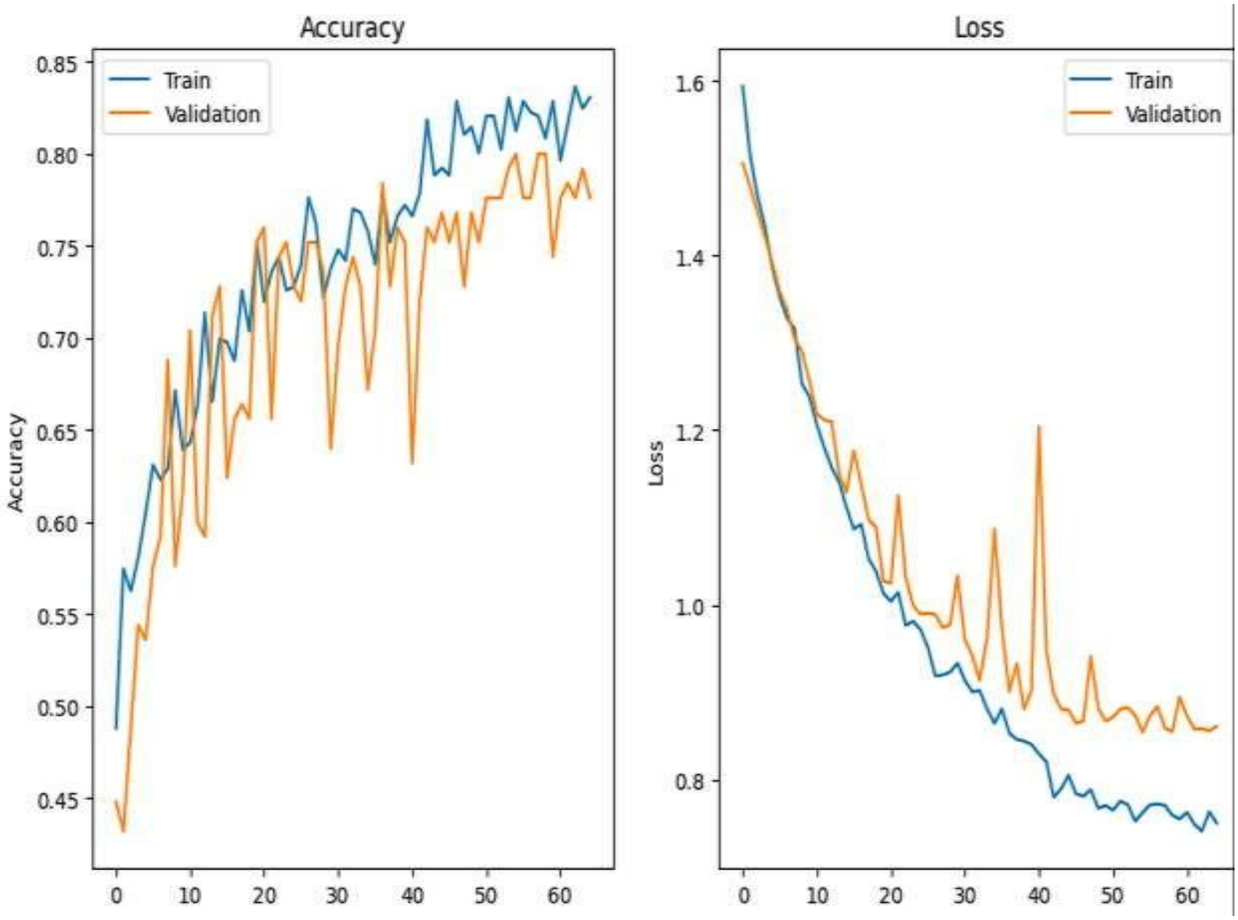

Fig3.Training and Validation Accuracy Across Epochs & Training and Validation Loss Across Epochs

Fig4. output for Abnormal activity detection

Fig5. output for Fall activity detection

Conclusion

The proposed study introduces a vision-based patient monitoring system integrating human pose estimation and Bidirectional GRU for fall detection, achieving an accuracy of 80%, precision of 83%, recall of 76%, and F1-score of 79%. The system effectively analyses temporal motion patterns using sequential pose data, improving detection performance compared to traditional and machine learning-based approaches. It enables non-intrusive monitoring using camera-based input and provides reliable real-time detection of abnormal activities.

The model requires moderate computational resources and is validated through performance metrics and graphical analysis. However, it is limited to video-based monitoring and may be affected by environmental conditions such as lighting and occlusion. Future work includes improving model robustness, integrating multi-modal data, and optimizing the system for real-time deployment. This system provides an efficient and scalable solution for enhancing patient safety in healthcare environments.

References

[1] Z. Qu et al., “Physics Sensor Based Deep Learning Fall Detection System,” 2024.
[2] J. D. Cárdenas et al., “Doppler Radar-Based Fall Detection Using Deep Learning,” 2023.
[3] T. Chen et al., “Wi-Fi Signal-Based Human Activity Recognition for Fall Detection,” 2024.
[4] N. T. Cam, “Fall Detection Using Human Pose Estimation and OpenPose Framework,” 2022.
[5] R. K. Chakravarthy and R. S. C. Reddy, “Machine Learning Approaches for Fall Detection Systems,” 2024.
[6] Google AI, “MediaPipe BlazePose: Real-Time Human Pose Estimation,” 2020.
[7] K. Cho et al., “Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation,” 2014.
[8] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” 2015.