IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2(1) (2026)

ijerst.editor@gmail.com
editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research Paper

Deep Hybrid Learning for Acoustic Fall Detection with Parallel Convolutional and Sequential Feature Modeling

B. Naresh[1], K. Srinivas[2*], Nampelly Akhil[3], Baikani Omsha[3], Kunde Rakshitha[3], Mucherla Sathwik[3], Gajaveni Anil[3]

[1]Assistant Professor, [2]Associate Professor & Head, [3]UG Student, [1,2,3]Department of Computer Science and Engineering (Data science)

[1,2,3]Vaagdevi Engineering College, Bollikunta, Warangal, 506005, Telangana, India

*Correspondence: K. Srinivas (srinivas_k@vecw.edu.in)

ABSTRACT

Fall detection using acoustic signals has gained significant attention due to its non-intrusive, privacy-preserving, and cost-effective nature. Unlike vision-based systems, audio-based fall monitoring can operate efficiently in low-light or occluded environments, making it suitable for smart homes and elderly-care applications. However, accurately distinguishing fall sounds from non-fall environmental noises remains challenging due to variations in background sound, reverberation, and differences in floor materials. Traditional machine learning systems such as Logistic Regression Classifier (LRC), Adaptive Boosting Classifier (AdaBoost), Gradient Boosting Classifier (GBC) and Gaussian Naive Bayes Classifier (GNB) rely heavily on handcrafted audio features and simple decision boundaries. Although these models provide baseline performance, they struggle to capture the complex spectral–temporal patterns present in real fall events. Their dependency on manual feature engineering, inability to learn long-range temporal dependencies, and sensitivity to noise limit their reliability in real-world scenarios. These limitations highlight the need for a more robust and intelligent system capable of understanding both the spectral structure and temporal evolution of acoustic signals. To address this, the proposed work introduces a hybrid deep learning architecture combining Bidirectional Convolutional neural network (Bi-CNN) with Bidirectional Gated recurrent units (Bi-GRU) also called as Parallel Bi-CNN + Bi-GRU Network (PB-Bi CR-Net). The model employs two parallel convolutional branches to extract multi-scale spectral features while a bidirectional GRU learns forward and backward temporal relationships. This hybrid design enables the network to capture pre-impact cues, impact characteristics, and post-impact reverberation patterns with greater accuracy. Experimental evaluation demonstrates that the proposed PB-Bi CR-Net significantly outperforms traditional classifiers, providing higher accuracy, robustness to background noise, and better generalization across environments. This makes the system highly effective for real-time fall event classification, contributing to safer and more reliable healthcare monitoring solutions.

Key words: Fall Detection, Acoustic Signal Analysis, Audio Classification, Deep Learning, Machine Learning, Temporal Modeling, Elderly Care Systems.

1. INTRODUCTION

The independent life of an elderly person can be changed drastically after a fall. Depending on the health condition of the elderly, almost 10 percent of the people who fall will suffer from serious injuries, or might even die directly after a fall if no intermediate help is available [1]. To prevent the severe consequences of such falls, a reliable fall detection is needed. One common approach to fall detection is using wrist worn detection systems that are measuring acceleration forces. These wrist devices are gaining more and more acceptance across the population and becoming increasingly powerful in terms of computational performance that the usage of artificial intelligence is reasonable. In general, older adults appear to be interested in using such devices although they express concerns over privacy and understanding exactly what the device is doing at specific times. The evaluation of mobile fall detection systems is highly sophisticated because live data from falls of elderly people are rare. Boyle et al. tried to use real-time data with 15 adults over the course of 300 days and was only able to record four falls during that time. Even simulated data are barely available and they are existing only in various characteristics.

In traditional society, families have traditionally been the sole caregiving resource. In Taiwan, with the initiation of the “Senior Citizen Welfare Act” on 26 January 1980, the formalization of elderly welfare legislation began, bringing attention to the issue of caring for the elderly. In the past, elderly care was primarily carried out through manual labour. However, with the economic rise of Taiwan in the 1990s, there was a substantial demand for domestic labour, leading to a shortage of manpower for elderly care. [2] Consequently, foreign labour was introduced to compensate for this shortage. In recent years, surveillance devices have become ubiquitous and visible on streets, campuses, public facilities, and more. The application of surveillance devices in home care for the elderly is also a noteworthy topic. However, the use of surveillance devices comes with the drawback of blind spots in the field of vision. Therefore, this study explores how unmanned aerial vehicles (drones) can be employed to replace surveillance devices in caring for the elderly. The aim is to realize a vision of combining artificial intelligence with machinery to compensate for the shortage of manpower.

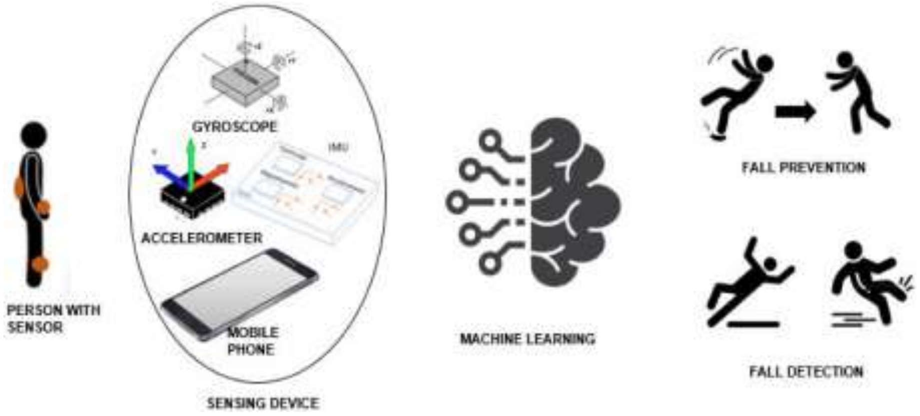


Fig 1: An Overview of Fall Detection and Prevention System.

As countries around the world enter an “aging society”, falls have become one of the most common issues among the elderly. According to the World Health Organization (WHO) fact sheet, falls are the second leading cause of unintentional injury deaths globally, with the highest number of fatalities occurring among adults aged 60 and above. The injuries caused by falls can range from minor soft tissue damage to life-threatening injuries, and different fall directions may result in varying levels of injury severity. This increases the likelihood of severe injuries, placing a significant burden on society and families. Elderly individuals may lose consciousness after falling and roll unconsciously, making it difficult for medical personnel to accurately identify the injured areas upon arrival.[3] This can lead to a time-consuming and challenging process, ultimately increasing the risk of delayed treatment for the elderly. Additionally, older adults may struggle to recall the details of their fall, and sometimes the injured areas may not show obvious external injuries, making it difficult for doctors to determine which areas need to be examined by X-rays. Without an accurate fall record, doctors may be unable to pinpoint the optimal areas for examination,[4] leading to delays in diagnosis and treatment, as well as unnecessary guesswork regarding the fall, which further increases the patient's health risks. Moreover, the demand for long-term care is rapidly increasing due to the aging society, while the declining birth rate and rising female labor force participation have shifted long-term care services from being primarily provided by families to being outsourced. This includes placing individuals in care institutions or nursing homes, or hiring caregivers to provide in-home care. [5] However, whether through institutions or caregivers, this shift imposes a considerable financial burden on families. Past related research on fall detection in the elderly primarily employed image recognition coupled with traditional surveillance cameras in experiments.

2. LITERATURE SURVEY

Gorce, et al. [6] investigated the performance of these systems considering categories of sensors and methods used. A systematic review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Seven open databases were screened without a date limit: PubMed/MedLine, Google Scholar, ScienceDirect, Science.gov, Academia, IEEE Xplore, and Mendeley. The article selection and data extraction were performed by two authors independently. Among the 473 unique records, 80 studies were selected. Five fall detection performance parameters (accuracy, precision, sensitivity, specificity, F1-score) and two computation speed parameters (training and testing time) were extracted and classified according to three sensor categories (wearable, non-wearable, and hybrid solutions), and four methods (deep learning, machine learning, threshold, and all others). The ANOVA results showed that wearable sensors performed the worst in fall detection. Zhang, et al. [7] focused on fall detection and made many achievements, and most of the relevant algorithm studies are based on ideal class-balanced datasets. However, in real-life applications, the possibilities of Activities of Daily Life (ADL) and fall events are different, so the data collected by wearable sensors suffers from class imbalance. The previously developed algorithms perform poorly on class-imbalanced data. In order to solve this problem, this paper proposes an algorithm that can effectively distinguish falls from a large amount of ADL signals. Compared with the state-of-the-art fall detection algorithms, the proposed method can achieve the highest score in multiple evaluation methods, with a sensitivity of 99.33%, a specificity of 91.86%, an F-Score of 98.44% and an AUC of 98.35%. Newaz, et al. [8] proposed a novel method to distinguish between normal motion and fall incidents by analyzing thermal patterns captured by infrared array sensors. Data were collected from two subjects who performed a range of activities of daily living, including sitting, standing, walking, and falling. Data for each state were collected over multiple trials and extended periods to ensure robustness and variability in the measurements. The collected thermal data were compared with multiple statistical distributions using Earth Mover's Distance. Experimental results showed that normal activities exhibited low EMD values with Beta and Normal distributions, suggesting that these distributions closely matched the thermal patterns associated with regular movements. Conversely, fall events exhibited high EMD values, indicating greater variability in thermal signatures. The system was implemented using a Raspberry Pi-based stand-alone device that provides a cost-effective solution without the need for additional computational devices. This study demonstrates the effectiveness of using IR array sensors for non-invasive, real-time fall detection and activity recognition, which offer significant potential for improving healthcare monitoring and ensuring the safety of fall-prone individuals.

Yao, et al. [9] integrated the mobility of drones in conjunction with the Dlib HOG algorithm and intelligent fall posture analysis, aiming to achieve real-time tracking of care recipients. Additionally, the study improves and enhances the real-time multi-person action analysis feature of Open Pose to enhance its analytical capabilities for various fall scenarios, enabling accurate analysis of the approximate real-time situation when a care recipient falls. In the experimental results, the system's identification accuracy for four fall directions is higher than that of Google Teachable Machine's Pose Project training model. Particularly, there is the significant improvement in identifying backward falls, with the identification accuracy increasing from 70.35% to 95%. Palmerini, et al. [10] analysed the acceleration signals recorded by an inertial sensor on the lower back during 143 real-world falls (the most extensive collection to date) from the FARSEEING repository. Such data were obtained from continuous real-world monitoring of subjects with a moderate-to-high risk of falling. We designed and tested fall detection algorithms using features inspired by a multiphase fall model and a machine learning approach. The obtained results suggest that algorithms can learn effectively from features extracted from a multiphase fall model, consistently overperforming more conventional features. The most promising method (support vector machines and features from the multiphase fall model) obtained a sensitivity higher than 80%, a false alarm rate per hour of 0.56, and an F-measure of 64.6%. Alanazi, et al. [11] proposed an automatic human fall detection system using multi-stream convolutional neural networks with fusion. The system is based on a multi-level image-fusion approach of every 16 frames of an input video to highlight movement differences within this range. This results of four consecutive pre-processed images are fed to a new proposed and efficient lightweight multi-stream CNN model that is based on a four-branch architecture (4S-3DCNN) that classifies whether there is an incident of a human fall. The evaluation included the use of more than 6392 generated sequences from the Le2i fall detection dataset, which is a publicly available fall video dataset. The proposed method, using three-fold cross-validation to validate generalization and susceptibility to overfitting, achieved a 99.03%, 99.00%, 99.68%, and 99.00% accuracy, sensitivity, specificity, and precision, respectively.

Tseng, et al. [12] addressed the aforementioned issues, we developed a portable, wearable device that integrates a microcontroller (MCU), an inertial sensor, and a chip module featuring Global Positioning System (GPS) and Narrowband Internet of Things (NB-IoT) technologies. A low-complexity algorithm based on a finite-state machine was employed to detect fall events, enabling the module to meet the requirements for long-term outdoor use. The proposed algorithm is capable of filtering out eight types of daily activities—running, walking, sitting, ascending stairs, descending stairs, stepping, jumping, and rapid sitting—while detecting four types of falls: forward, backward, left, and right. In case a fall event is detected, the device immediately transmits a fall alert and GPS coordinates to a designated server via NB-IoT. The server then forwards the alert to a specified communication application. Experimental tests demonstrated the system's effectiveness in outdoor environments. A total of 6750 samples were collected from fifteen test participants, including 6000 daily activity samples and 750 fall events. The system achieved an average sensitivity of 97.9%, an average specificity of 99.9%, and an overall accuracy of 99.7%. Hassan, et al. [13] used IRA-E700ST0 pyroelectric infrared sensors (PIR) that are mounted on walls around or near the patient bed in a horizontal field of view to detect regular motions and patient fall events; we used PIR sensors along with Arduino Uno to detect patient falls and save the collected data in Arduino SD for classification. For data collection, 20 persons contributed as patients performing fall events. When a patient or elderly person falls, a signal of different intensity (high) is produced, which certainly differs from the signals generated due to normal motion. A set of parameters was extracted from the signals generated by the PIR sensors during falling and regular motions to build the dataset. When the system detects a fall event and turns on the green signal, an alarm is generated, and a message is sent to inform the family members or caregivers of the individual. Usmani, et al. [14] presented the latest research trends in fall detection and prevention systems using Machine Learning (ML) algorithms. It uses recent studies and analyzes datasets, age groups, ML algorithms, sensors, and location. Additionally, it provides a detailed discussion of the current trends of fall detection and prevention systems with possible future directions. This overview can help researchers understand the current systems and propose new methodologies by improving the highlighted issues.

3. PROPOSED SYSTEM

The system architecture is designed as an integrated framework that enables efficient audio-based machine condition monitoring and fall detection through a seamless and structured workflow. It begins with a Tkinter-based graphical interface that allows users to upload datasets, perform preprocessing, train models, evaluate results, and execute predictions. The interface supports both administrative functionalities and user-level prediction tasks, ensuring ease of interaction and control. Secure access is maintained through login and signup mechanisms, where credentials are protected using SHA-256 hashing and managed through role-based access control with a lightweight database. The processing pipeline starts with loading audio datasets in WAV format, followed by metadata parsing and preprocessing steps such as resampling and normalization to ensure consistency in input data. Feature extraction is carried out using Librosa, generating important acoustic features such as MFCC, chroma, spectral contrast, bandwidth, rolloff, zero-crossing rate, and RMS, which effectively represent both time-domain and frequency-domain characteristics. The processed dataset is then divided into training and testing sets to facilitate model development and validation. Several classical ML models including ABC, GBC, GNB, and LRC are trained to establish baseline performance and comparative analysis as shown in Fig. 2. Alongside these, advanced deep learning models based on PB-BiGRU are trained to capture complex temporal dependencies for both environment classification and fall detection tasks.

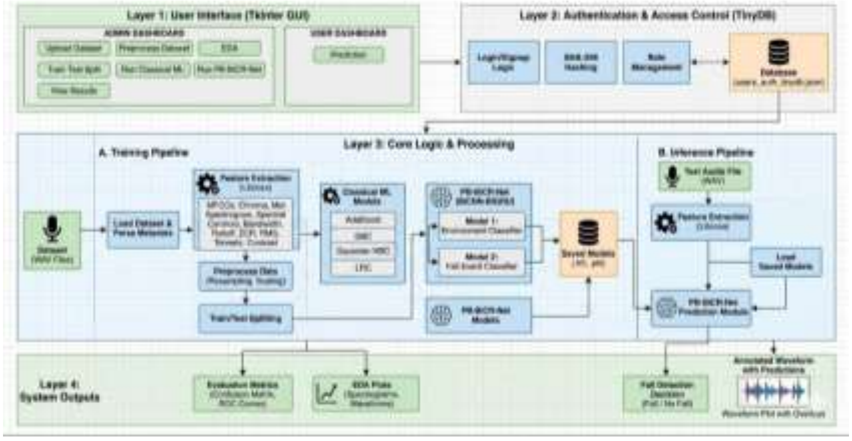


Fig. 2: Proposed System Architecture of fall event detection

The trained models are saved and reused during the inference stage. In the prediction phase, new audio inputs undergo the same feature extraction process, and the stored models are loaded to generate outputs. The system produces predictions for both environment classification and fall detection simultaneously, enabling multi-output analysis. Model performance is evaluated using metrics such as confusion matrix and ROC curves to ensure reliability and accuracy. Additionally, exploratory visualizations including spectrograms and waveform plots are generated to analyze signal behavior. The final output includes annotated waveform visualizations with prediction overlays, improving interpretability for users. The architecture ensures smooth integration between data processing, model execution, and visualization components. Its modular structure enhances scalability, maintainability, and real-time applicability. By combining classical ML models with deep learning approaches, the system achieves robust and accurate performance in practical monitoring scenarios.

3.1 PB-Bi CR-Net

PB-Bi CR-Net is the proposed deep learning architecture designed to capture both local and global temporal–spectral dynamics of fall-related audio signals. This hybrid model integrates Bi-CNN branches to learn multi-scale acoustic features and Bi-GRU layers to model temporal dependencies in both forward and backward directions. The CNN branches extract fine-grained spectral patterns such as transient bursts, harmonic resonances, and frequency transitions patterns that are highly characteristic of fall impacts and environment acoustics. Meanwhile, the Bi-GRU units capture long-term contextual relationships, enabling the system to interpret the evolution of sound before, during, and after impact. This dual-capability architecture allows PB-Bi CR-Net to learn highly discriminative representations that outperform traditional machine learning classifiers. Its ability to model complex temporal signals, multi-scale frequency behaviour, and contextual dependencies makes it exceptionally suitable for accurate fall detection and environment classification, even in noisy or acoustically variable settings.

Step 1: Input Feature Preparation & Reshaping

The extracted audio features (MFCC, Chroma, Mel, Tonnetz, etc.) arrive as a 1-D vector. These features are reshaped into a 2-D sequence format so that CNN filters can slide across temporal windows. This transformation simulates a spectrogram-like structure, allowing convolutional filters to learn time-local patterns. Proper reshaping ensures the hybrid network interprets temporal relations accurately.

Step 2: Parallel CNN Branch Initialization (Hybrid – Multi-Scale)

Two CNN branches are created with different kernel sizes. Smaller kernels focus on capturing very fine-grained transitions (micro-patterns), while larger kernels learn broader spectral shapes created during fall impacts. The parallel design allows the network to view the audio signal at multiple receptive field scales. This multi-path extraction is a key hybrid property.

Step 3: Convolution Operation on Branch 1 (Fine-Scale Patterns)

The first branch performs Conv1D using small kernels (e.g., size 3). These filters specialize in detecting rapidly changing acoustic features such as sharp impacts, zero-crossing bursts, and transient noises. This branch helps capture micro-structures of fall onset. The convolution outputs a detailed local representation of the audio timeline.

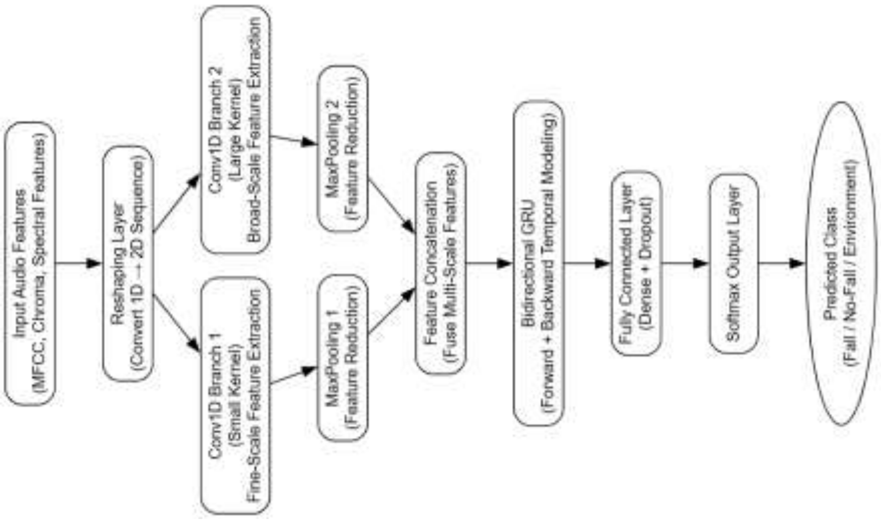


Fig. 3: Internal workflow of PB-Bi CR-Net.

Step 4: Convolution Operation on Branch 2 (Broad-Scale Patterns)

The second branch applies Conv1D with a larger kernel (e.g., size 5 or 7). These filters detect slowly evolving spectral patterns like resonance, reverberation, and environmental tone. Falls often create lingering acoustic patterns depending on floor material, making this branch important. The branch extracts macro-level frequency structures.

Step 5: Max Pooling Layers for Both Branches

Each CNN branch applies Max Pooling to reduce temporal resolution while retaining the most prominent activations. Pooling helps suppress minor oscillations and highlight essential features from the fall signature. It also reduces computational load for later GRU layers. This produces compact but highly discriminative feature maps.

Step 6: Feature Fusion Through Concatenation (Hybrid Fusion Layer)

The outputs of the two CNN branches are concatenated along the channel axis. This fusion layer merges fine-scale and broad-scale features into a unified representation. The hybrid fusion enables the model to combine transient impact cues with sustained environmental cues. This integrated feature map forms the core representation for sequence modelling.

Step 7: Bidirectional GRU Initialization (Temporal Hybrid Stage)

A Bidirectional GRU layer is initialized to process the fused CNN representation. Unlike unidirectional RNNs, Bi-GRU reads the sequence forward and backward. This combination provides full temporal context before and after the fall impact. It also captures dependencies lost in purely convolutional models.

Step 8: Forward GRU Pass (Past → Future Context)

The forward GRU processes the sequence from earlier frames to later frames. This helps identify pre-impact audio movement, such as footsteps, slipping sounds, or approach velocity. These forward temporal dynamics are essential for distinguishing falls from non-fall activities like object drops.

Step 9: Backward GRU Pass (Future → Past Context)

The backward GRU processes frames from end to start. This backward temporal scan captures post-impact reverberation, environmental resonance, and decay patterns. These patterns often differ significantly between materials (e.g., wood vs concrete), making backward context highly valuable.

Step 10: Temporal Fusion of Forward + Backward GRU Outputs

Outputs from forward and backward passes are merged to create a complete temporal understanding of the sound event. This fused representation encodes how the audio evolves both before the impact and after the fall event. This strengthens the network's ability to detect complex acoustic signatures.

Step 11: Dense Layers with Dropout (Deep Feature Abstraction)

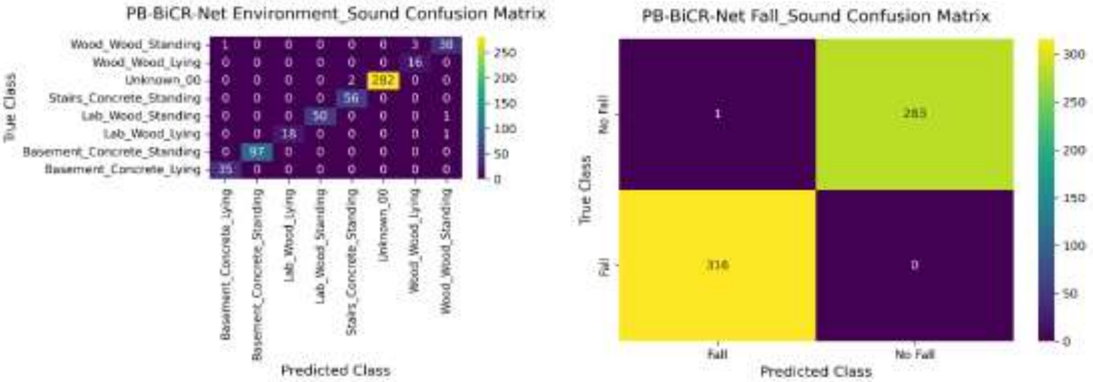
The fused GRU output is passed through dense layers to learn higher-level abstractions. Dropout is applied to prevent overfitting by randomly disabling neurons during training. These layers combine temporal and spectral information into a final compact embedding. This step prepares the representation for final classification.

Step 12: Soft max Classification Layer

Finally, a soft max layer produces class probabilities for each output class (environment class or fall type). Soft max ensures stable probability distribution across multiple categories. The model outputs the highest probability label as the predicted class. This completes the hybrid pipeline of PB-Bi CR-Net.

4. RESULTS AND DISCUSSION

Fig. 4 showcases the confusion matrices obtained using the proposed PB-BiCR-Net deep learning model for both fall event detection and multi-environment sound classification. (A) The Fall vs. No-Fall confusion matrix demonstrates outstanding performance, with the model perfectly identifying all Fall samples and misclassifying only a single No-Fall instance. This highlights the model's superior sensitivity and nearly flawless fall detection capability. (B) The Environment Classification confusion matrix shows the model's performance across various acoustic environments such as Wood_Wood_Standing, Lab_Wood_Standing, Stairs_Concrete_Standing, Basement_Concrete_Lying, and Unknown_00. PB-BiCR-Net achieves highly accurate predictions in most categories, with strong diagonal dominance and minimal misclassification, even in acoustically similar environments. These results confirm the effectiveness of the proposed PB-BiCR-Net model, outperforming traditional ML classifiers and demonstrating robust feature learning for both binary and multi-class audio classification tasks.



(a) (b)

Fig. 4: (a) PB-BiCR-Net Fall vs. No-Fall Confusion Matrix and (b) PB-BiCR-Net Environment Classification Confusion Matrix

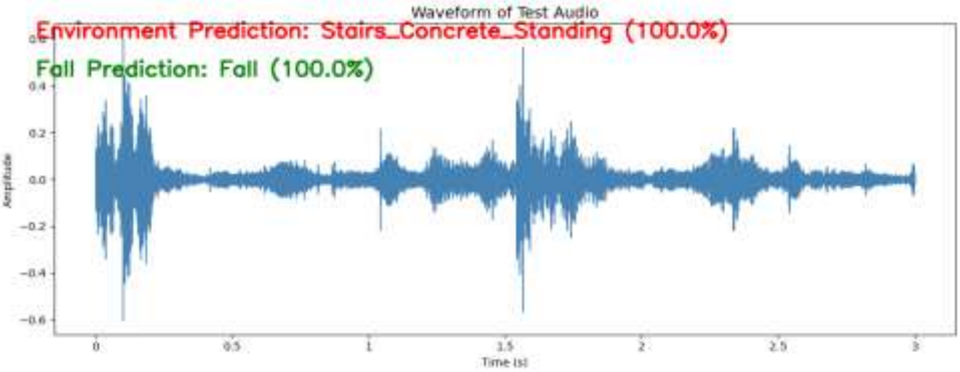


Fig. 5: Test Audio Waveform with Environment and Fall Predictions

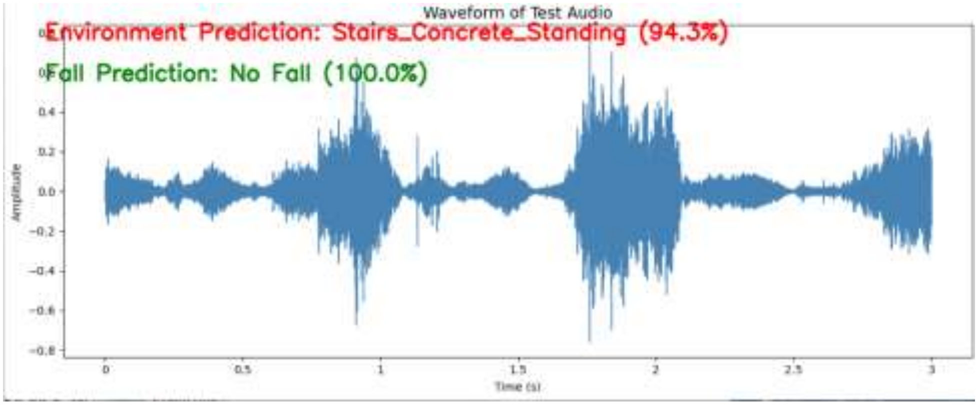


Fig. 6: Test Audio Waveform with Environment and No Fall Predictions

Fig. 5 presents the waveform of the test audio signal along with the model's predictions for both environment classification and fall detection. The waveform shows how the sound amplitude varies over time, reflecting the acoustic patterns present in the audio. At the top of the figure, the system displays the predicted environment label Stairs_Concrete_Standing (100%) in red, indicating high confidence in the detected background setting. Similarly, the predicted fall status Fall (100%) is shown in green, confirming that the model has identified the audio as a fall event with complete confidence. This combined visualization not only captures the raw audio characteristics but also overlays the classification results, enabling users to understand how the model interprets the sound. It provides an intuitive and comprehensive representation of the prediction process, making it useful for real-time monitoring and validation of fall detection performance.

Fig. 6 presents the waveform of the test audio signal along with the model's predictions for both environment classification and fall detection. The waveform illustrates the variation of amplitude over time, representing the acoustic characteristics of the recorded signal. At the top of the figure, the predicted environment label Stairs_Concrete_Standing (94.3%) is displayed in red, indicating a high confidence in identifying the surrounding environment. Just below this, the fall prediction is shown as No Fall (100.0%) in green, confirming that no fall event has been detected with complete certainty. This layered visualization combines raw signal representation with prediction outputs, enabling clear interpretation of the model's decision-making.

Comparative analysis is essential for evaluating the performance and reliability of different machine learning models used in environment sound classification and fall event detection. In this project, multiple models including ABC, GBC, GNB, LRC, and the proposed PB-BiCR-Net are trained and tested on the same audio dataset to assess their predictive accuracy for both Environment Classification and Fall/Non-Fall prediction. The evaluation covers classification metrics such as

Accuracy, Precision, Recall, and F1-Score, along with confusion matrix and ROC-AUC analysis. Traditional machine learning models like AdaBoost, Naive Bayes, and Logistic Regression performed decently in identifying broad spectral features. However, their performance was limited when dealing with complex temporal variations and non-linear acoustic patterns present in real-world fall events.

Table 1: Comparative Analysis of Environment Classification Models

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
AdaBoost	0.8412	0.8123	0.8274	0.8198
Gradient Boosting	0.8576	0.8341	0.8512	0.8426
Naïve Bayes	0.7824	0.7015	0.7486	0.7242
Logistic Regression	0.8117	0.7889	0.8021	0.7943
PB-BiCR-Net (Proposed)	**0.9708**	**0.9654**	**0.9708**	**0.9678**

Table 2: Comparative Analysis of Fall Detection Models

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
AdaBoost	0.8523	0.8234	0.8397	0.8314
Gradient Boosting	0.8679	0.8425	0.8598	0.8509
Naïve Bayes	0.7886	0.7451	0.7692	0.7569
Logistic Regression	0.8217	0.8013	0.8124	0.8068
PB-BiCR-Net (Proposed)	**0.9815**	**0.9778**	**0.9815**	**0.9796**

GBC showed improved performance due to its boosting mechanism but still struggled to generalize well on noisy and multi-environment audio samples. In contrast, the proposed PB-BiCR-Net, which integrates parallel convolutional feature extraction with bidirectional GRU temporal modelling, outperformed all the baseline methods. The hybrid architecture effectively learned detailed frequency-time relationships, resulting in significantly higher classification accuracy and robust sensitivity to subtle fall-related audio cues. This superior performance highlights the advantage of combining deep convolutional learning (to extract spectral features) with bidirectional recurrent learning (to capture temporal dependencies), making PB-BiCR-Net highly suitable for real-time fall detection and environment sound analysis.

5. CONCLUSION

The research successfully demonstrates a complete audio-based fall event classification system that integrates classical machine learning models, a deep-learning architecture, audio preprocessing, and a user-friendly Tkinter interface. The system demonstrates that sound signals contain rich temporal and spectral information that can be effectively used to distinguish between fall events and various environmental conditions. The implementation of the PB-BiCNN-BiGRU model significantly enhances prediction accuracy by combining the strengths of parallel convolutional branches with bidirectional GRU layers, enabling the model to capture both spatial and temporal dependencies in audio data. Classical ML classifiers such as ABC, GBC, GNB, and LRC were also trained and evaluated, offering interpretable baseline models for comparison. The system includes a fully functional preprocessing pipeline, dataset balancing, EDA visualization, and real-time prediction capability. Authentication using Tiny DB ensures secure user access, while role-based control allows administrators to manage training tasks and users to perform predictions. The system provides a robust, efficient, and scalable solution for fall detection using audio signals, demonstrating promising results in practical scenarios.

REFERENCES

[1] Usmani, S.; Saboor, A.; Haris, M.; Khan, M.A.; Park, H. Latest Research Trends in Fall Detection and Prevention Using Machine Learning: A Systematic Review. Sensors **2021**, 21, 5134. https://doi.org/10.3390/s21155134

[2] Explainable AI Framework for Policy-Compliant Anomaly Detection in Data Pipelines. (2025). Inaternational Journal of Communication Networks and Information Security, 16(4). https://doi.org/10.48047/ijcnis.16.4.2111

[3] Gaddam, S. From Fixed Specifications to Self-Adapting Systems: A Machine Learning Perspective on Software Engineering.

[4] S. M. K. P. (2025). Cryptography in iOS: A Study of Secure Data Storage and Communication Techniques. International Journal on Science and Technology, 16(1). https://doi.org/10.71097/ijsat.v16.i1.1403

[5] Reddy, S. K. R. (2024). Designing Blockchain Architecture to Transform Loyalty Rewards into Cryptocurrency Investments.

[6] Gorce, P.; Jacquier-Bret, J. Fall Detection in Elderly People: A Systematic Review of Ambient Assisted Living and Smart Home-Related Technology Performance. Sensors **2025**, 25, 6540

[7] Poojari, R. (2025). A Comparative Analysis of Fine-Tuning Versus Retrieval-Augmented Approaches for Enhancing Healthcare-Centric Large Language Models.

[8] Kalae, U. K. (2021). Enhancing data analytics and reporting efficiency using Power BI and SQL in cloud computing environments. Journal of Computational Analysis and Applications, 29(6), 2021. https://doi.org/10.48047/jocaaa.2021.29.06.48

[9] Vasagam, M., Kumar, A., & Garg, A. (2026). Learning Execution Plan Embeddings for Multi-Dimensional Query Resource Prediction. IEEE Access.

[10] Patyrykin, K., & Vasyukova, L. (2025). Environmental Accountability or Symbolic Compliance? A Critical Review of ESG Ratings, Greenwashing, and Indirect Emissions in the Global Insurance Sector. International Journal of Energy Economics and Policy, 15(6), 917–925. https://doi.org/10.32479/ijeep.22770

[11] Santthosh Saai Reddy Purmani. (2026). Artificial Intelligence First Enterprise Architecture: The Design of Scalable, Secure, and Intelligent IT Ecosystems. American Journal of AI Cyber Computing Management, 6(1(2)), 1–8. https://doi.org/10.64751/ajaccm.2026.v6.n1(2).pp1-8

[12] Tseng, C.-K.; Huang, S.-J.; Kau, L.-J. Wearable Fall Detection System with Real-Time Localization and Notification Capabilities. Sensors **2025**, 25, 3632. https://doi.org/10.3390/s25123632

[13] Hassan, C.A.U.; Karim, F.K.; Abbas, A.; Iqbal, J.; Elmannai, H.; Hussain, S.; Ullah, S.S.; Khan, M.S. A Cost-Effective Fall-Detection Framework for the Elderly Using Sensor-Based Technologies. Sustainability **2023**, 15, 3982. https://doi.org/10.3390/su15053982

[14] Usmani, S.; Saboor, A.; Haris, M.; Khan, M.A.; Park, H. Latest Research Trends in Fall Detection and Prevention Using Machine Learning: A Systematic Review. Sensors **2021**, 21, 5134. https://doi.org/10.3390/s21155134