International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 6, Nov-Dec-2026, ISSN (Online): 2395-566X

AI for Human Activity Recognition Using Sensor Data

Kandhare Supriya Dilip
Sahayadri Valley College of Engineering and Technology Savitribai Phule Pune University (SPPU), Pune, Maharashtra, India

Abstract — Human Activity Recognition (HAR) is the process of identifying a person's physical activity from signals collected by sensors. This paper presents a proposed artificial intelligence framework for recognizing common human activities using inertial sensor data from smartphones or wearable devices. The system uses a tri-axial accelerometer and gyroscope to capture body motion, divides continuous signals into overlapping time windows, extracts temporal and statistical representations, and classifies each window using a hybrid one-dimensional Convolutional Neural Network and Long Short-Term Memory (1D-CNN–LSTM) model. The UCI Human Activity Recognition Using Smartphones dataset is proposed as the primary benchmark because it contains recordings from 30 volunteers performing six activities and provides accelerometer and gyroscope measurements sampled at 50 Hz. The paper describes data preparation, model design, training strategy, evaluation metrics, and deployment considerations. Since no new experiment or model execution was supplied for this draft, numerical performance is not claimed; a reproducible evaluation protocol is provided instead. The proposed approach aims to support applications such as fitness monitoring, context-aware mobile services, and assisted-living systems while considering privacy, user variability, and computational constraints.

Keywords— Human Activity Recognition, Artificial Intelligence, Accelerometer, Gyroscope, Sensor Data, Deep Learning, CNN-LSTM.

I. INTRODUCTION

The widespread availability of smartphones, smartwatches, and wearable Internet of Things devices has made continuous motion sensing practical. These devices contain inertial measurement sensors, especially accelerometers and gyroscopes, that measure linear acceleration and angular velocity along three axes. Machine-learning methods can interpret these time-series signals and infer activities such as walking, sitting, standing, climbing stairs, and lying.

Human Activity Recognition is useful in health and wellness monitoring, physical activity tracking, elderly-care support, rehabilitation, and context-aware computing. Manual observation is time-consuming and cannot provide continuous monitoring. Sensor-based recognition offers an automated alternative, but real-world signals vary with device placement, orientation, individual movement patterns, sensor noise, and transitions between activities.

This paper proposes an AI-based HAR pipeline that combines local feature learning through a one-dimensional convolutional neural network with temporal sequence learning through a Long Short-Term Memory network. The main objectives are: (1) to describe a reproducible sensor-data preparation pipeline; (2) to design a CNN–LSTM classifier for windowed inertial signals; (3) to define subject-independent evaluation using established metrics; and (4) to discuss practical limitations and privacy considerations.

II. RELATED WORK

Early smartphone-based HAR research demonstrated that accelerometer and gyroscope signals could distinguish common daily activities. The UCI benchmark introduced by Anguita and colleagues contains six activity classes and a subject-separated training and test partition. Its data representation includes time- and frequency-domain features, making it a widely used baseline for classification studies.

Traditional approaches commonly calculate handcrafted features such as mean, standard deviation, signal magnitude, energy, and frequency-domain coefficients before applying classifiers such as support vector machines, decision trees, or random forests. These methods can be computationally efficient, but feature design may require domain expertise and may not capture all temporal patterns.

Deep-learning approaches learn representations directly from sensor sequences. Convolutional networks can identify short-duration patterns in multi-axis signals, while recurrent models such as LSTM networks can represent longer temporal dependencies. Research on wearable activity recognition has explored CNN and recurrent architectures, and survey work highlights the importance of preprocessing, sensor placement, feature learning, and evaluation methodology. The proposed work uses these findings to motivate a compact hybrid architecture rather than claiming a new state-of-the-art result.

© 2026 IJSRET

III. PROPOSED SYSTEM

The proposed system accepts synchronized accelerometer and gyroscope readings as input and returns one activity label for each analysis window. The overall stages are sensor acquisition, signal cleaning, window segmentation, normalization, feature learning, classification, and output smoothing.

1. Sensor Acquisition

A smartphone or wearable device provides six inertial channels: acceleration along the x, y, and z axes (a_x, a_y, a_z) and angular velocity along the x, y, and z axes (g_x, g_y, g_z). If a device provides additional sensors, such as a magnetometer, those channels may be studied separately. Sensor timestamps should be retained so that irregular sampling and missing readings can be detected.

2. Preprocessing and Windowing

Raw measurements are checked for missing values, duplicated timestamps, and outliers. Signals are resampled to a consistent sampling rate when necessary, and each channel is normalized using statistics calculated from the training set only. A sliding window converts the continuous stream into fixed-length samples. For the UCI HAR benchmark, the published protocol uses 128 readings per window at 50 Hz (approximately 2.56 seconds) with 50% overlap. For a new device or dataset, window length should be selected using validation data and the intended response latency.

Each window is represented as a matrix $X \in R^{(T\times6)}$, where T is the number of time steps and six denotes the accelerometer and gyroscope axes. Labels are assigned from the annotated activity associated with the window. Windows spanning an activity transition should either be excluded under a clearly stated rule or assigned using a documented majority-label policy.

3. CNN–LSTM Model

The proposed model contains a one-dimensional convolutional block followed by an LSTM layer and a dense classification head. The convolutional block learns local movement patterns across the six sensor channels. Batch normalization and ReLU activation can stabilize training, while max pooling reduces sequence length. The LSTM layer models temporal dependencies across the resulting feature sequence. A dropout layer is used to reduce overfitting, and a softmax output layer produces class probabilities for the six target activities.

For an input window X, the convolutional representation is $H = CNN(X)$, the temporal representation is $z = LSTM(H)$, and the predicted class is $y_hat = argmax(softmax(Wz + b))$. The network is trained using categorical cross-entropy and an optimizer such as Adam. Hyperparameters—including filter count, kernel size, LSTM units, dropout rate, learning rate, and batch size—must be selected on the validation set rather than on the test set.

IV. DATASET AND EXPERIMENTAL METHOD

1. Benchmark Dataset

The primary recommended benchmark is the UCI Human Activity Recognition Using Smartphones dataset. It contains 10,299 labeled windows collected from 30 volunteers aged 19–48 while carrying a waist-mounted smartphone. The six activities are walking, walking upstairs, walking downstairs, sitting, standing, and laying. The smartphone's tri-axial accelerometer and gyroscope were sampled at 50 Hz. The dataset provides a predefined subject-wise split, with approximately 70% of participants used for training and 30% for testing. This split should be preserved to reduce identity leakage between training and test participants.

A second optional dataset is the WISDM Smartphone and Smartwatch Activity and Biometrics Dataset. It includes accelerometer and gyroscope streams from both a smartphone and smartwatch for 51 participants performing 18 activities, sampled at 20 Hz. It can be used in future work to test cross-device and larger activity-set generalization; its sampling rate and sensor arrangement require a separate preprocessing configuration.

2.Training Procedure

The training partition is divided into training and validation subsets by participant wherever feasible. Normalization parameters are estimated only from the training subset and then applied unchanged to validation and test data. The CNN–LSTM is trained with early stopping based on validation loss. Class weights may be used if the activity distribution is imbalanced. A fixed random seed, software versions, model hyperparameters, and preprocessing configuration should be recorded to support reproducibility.

3. Evaluation Metrics

Performance should be reported using overall accuracy, macro-averaged precision, macro-averaged recall, macro F1-score, and a confusion matrix. Macro F1 is important because it gives each activity class equal weight. Per-class recall should be

© 2026 IJSRET

inspected to identify confusion between similar activities, particularly sitting versus standing and walking upstairs versus walking downstairs. For deployment, inference latency, model size, and energy consumption should also be measured on the target device.

4. Baseline Comparison

The proposed CNN–LSTM should be compared with at least one conventional baseline, such as a Random Forest or Support Vector Machine trained on the dataset's supplied feature vectors, and a simpler 1D-CNN model trained on the same raw-window split. All models must use identical subject-level partitions. Report mean and standard deviation over repeated training runs if computational resources permit.

V. SYSTEM ARCHITECTURE

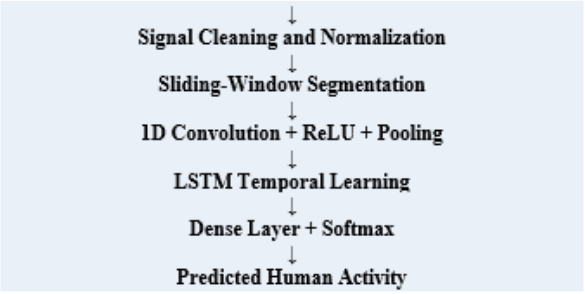

Fig. 1. Block diagram of the proposed sensor-based human activity recognition system.

VI. RESULTS AND DISCUSSION

No experimental run, trained model, or measured prediction output was provided when this manuscript was prepared. Therefore, this section does not report fabricated accuracy, precision, recall, F1-score, or execution-time values. After implementation, the results should be inserted in Table I and discussed using the held-out subject test set.

Table 1. Experimental results to be completed after model evaluation.

Model	Accuracy	Macro Precision	Macro Recall	Macro F1
Random Forest	To measure	To measure	To measure	To measure
1D-CNN	To measure	To measure	To measure	To measure
Proposed CNN–LSTM	To measure	To measure	To measure	To measure

The confusion matrix should be used to determine which activity pairs are difficult to distinguish. If the model performs well on the predefined test split but degrades on participants or devices not represented in training, this may indicate subject or device dependence. Additional evaluation should therefore include participant-wise cross-validation and, where possible, a separate dataset such as WISDM. The latency and memory footprint should be measured before claiming suitability for real-time or edge deployment.

Applications and Limitations

Potential applications include activity summaries for fitness applications, context-aware smartphone interfaces, rehabilitation progress monitoring, and supportive monitoring for older adults. These applications should not treat activity predictions as medical diagnoses or as a substitute for clinical assessment.

Important limitations include variation in phone placement, differences in sensor calibration, individual movement styles, class imbalance, and ambiguity during transitions. A model trained on waist-mounted smartphone data may not generalize to a wrist-worn device without adaptation. Sensor streams can also reveal behavioral patterns; therefore, data minimization, informed consent, secure storage, and on-device inference should be considered wherever possible.

VII. CONCLUSION

This paper presented a proposed AI framework for human activity recognition using accelerometer and gyroscope data. The framework combines preprocessing and sliding-window segmentation with a 1D-CNN–LSTM classifier to learn spatially local motion patterns and temporal dependencies. The UCI HAR dataset provides a reproducible starting point, while WISDM offers a possible extension to multiple devices and a wider range of activities. The paper deliberately defines an evaluation protocol without inventing experimental results. Future work should implement the model, compare it with conventional and deep-learning baselines, measure subject-independent performance, and evaluate efficiency on a real smartphone or wearable device.

REFERENCES

1. D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, “A Public Domain Dataset for Human Activity Recognition Using Smartphones,” in Proc. 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN), 2013.

© 2026 IJSRET

2. J. L. Reyes-Ortiz, D. Anguita, A. Ghio, L. Oneto, and X. Parra, “Human Activity Recognition Using Smartphones,” UCI Machine Learning Repository, 2013, doi: 10.24432/C54S4K.
3. A. M. Khan, Y.-K. Lee, S. Y. Lee, and T.-S. Kim, “A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer,” IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 5, pp. 1166–1172, 2010.
4. F. J. Ordóñez and D. Roggen, “Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition,” Sensors, vol. 16, no. 1, article 115, 2016, doi: 10.3390/s16010115.
5. O. D. Lara and M. A. Labrador, “A survey on human activity recognition using wearable sensors,” IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1192–1209, 2013, doi: 10.1109/SURV.2012.110112.00192.
6. G. M. Weiss, K. Yoneda, and T. Hayajneh, “Smartphone and smartwatch-based activity recognition using deep learning,” in Proc. IEEE International Conference on Healthcare Informatics, 2019. (Consult the WISDM dataset documentation for the exact dataset version and citation details before submission.)
7. G. M. Weiss, “WISDM Smartphone and Smartwatch Activity and Biometrics Dataset,” UCI Machine Learning Repository, 2019, doi: 10.24432/C5HK59.
8. A. B. N. Nweke, Y. W. Teh, M. A. Al-garadi, and U. R. Alo, “Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges,” Expert Systems with Applications, vol. 105, pp. 233–261, 2018.

© 2026 IJSRET