

## Article

# Wearable IoT and Artificial Intelligence Techniques for Leveraging the Human Activity Analysis

Lina Sheker<sup>1</sup>, Vishwanath Petli<sup>2</sup>, K. Satish Reddy<sup>3</sup>

<sup>1</sup> Department of electrical and computer engineering, Altinbaş University, Istanbul, 34000, Turkey

<sup>2</sup> D Department Electronics and Communication Engineering, H.K.E. Society's Sir M Visvesvaraya College of Engineering, Raichur, India

<sup>3</sup> Department of Electronics and Communication Engineering, Hyderabad Institute of Technology and Management, Hyderabad, India

\*Corresponding author: Lina Sheker (email: 2037203@ogr.altinbas.edu.tr)

Received 05 March 2024; Accepted 05 March 2024; Published 15 June 2024

**Abstract:** The integration of wearable Internet of Things (IoT) devices with advanced Artificial Intelligence (AI) techniques has significantly transformed human activity analysis (HAA). It enables precise, real-time monitoring of diverse physical behaviors. This research delves into the design of an efficient and accurate human activity recognition (HAR) system utilizing the WISDM dataset. The dataset comprises accelerometer and gyroscope data collected from wearable devices during various activities, offering a robust foundation for HAR. Despite advancements, existing systems face challenges in accurately capturing the spatial and temporal dependencies in time-series data, which often leads to suboptimal classification performance. To address this, a hybrid deep learning framework is proposed, combining Convolutional Neural Networks (CNN) for spatial feature extraction and Long Short-Term Memory (LSTM) networks for temporal sequence modeling. This combination ensures better learning by using CNNs for spatial features and LSTMs for time-based patterns, making activity analysis more accurate. The proposed model attains a remarkable classification accuracy of 99.4%, underscoring its efficacy in distinguishing between various activities in real-time scenarios. Key performance metrics, including Precision (99%), Recall (98.6%), Specificity (99.3%), and F1-Score (99.2%), are systematically evaluated, validating the model's robustness, reliability, and practical applicability. Experimental results reveal the model's consistent and exceptional performance, highlighting its superiority over existing state-of-the-art approaches. The findings underscore the practical viability of deploying such intelligent systems in real-world environments, representing a notable progression in the evolution of human activity analysis.

**Keywords:** Wearable IoT, Human Activity Recognition (HAR), WISDM Dataset, CNN, LSTM, Deep Learning Framework.

## 1. Introduction

Recently, the incorporation of wearable IOT devices with AI methods has potentially transformed the landscape of human activity analysis. The advancements in wearable IoT technology, combined with the power of AI algorithms, have opened up new avenues for real-time monitoring, personalized healthcare, fitness tracking, and even the enhancement of daily life [1-4]. Wearable IoT devices such as smartwatches, fitness trackers, and smart clothing are capable of continuously collecting massive volume of information on various human activities, including physical movements, vital signs, and environmental factors. This data, once processed and analyzed using AI techniques, can provide a wealth of insights into an individual's health, lifestyle, and overall well-being [5].

Wearable IoT devices typically incorporate a range of sensors, like accelerometers, gyroscopes, heart rate monitors, GPS units, and thermometers, to track diverse aspects of human behavior. These sensors collect data on activities such as walking, running, sitting, sleeping, and even more complex behaviors like cycling or specific sports-related movements. This data is then transmitted to cloud platforms or local devices where it can be stored, processed, and analyzed [6-8]. The growing availability of high-quality sensors and advanced communication technologies, such as 5G and Bluetooth, enables these devices to operate seamlessly in real-time, offering continuous monitoring without the need for user intervention.

The true ability of wearable IoT devices lies in the integration of AI techniques particularly machine learning (ML) [9-10] and deep learning (DL), can process and assess the vast amount of information generated by these devices to uncover patterns, predict future activity, and offer actionable insights. ML models, like supervised and unsupervised learning, can be trained to classify varied types of human activity. DL algorithms like CNN and Recurrent Neural Networks (RNNs), can go beyond simple classification, enabling more sophisticated tasks such as activity recognition, gesture detection, and even the prediction of future health events or physical performance [11-12].

Despite the significant advancements in wearable IoT devices and AI, there are several challenges associated with existing algorithms for human activity analysis. One major challenge is the complexity of processing vast amount of sensor data in real-time. The sheer amount of information generated by wearable devices can overwhelm traditional machine learning models, requiring advanced computational techniques to process efficiently. Additionally, the variability of human activities and individual behavior adds complexity to activity recognition, often leading to issues with accuracy and generalization. Another challenge is the limited ability of existing models to handle noisy or incomplete data. Furthermore, the integration of heterogeneous sensor data from different IoT devices and the need for continuous monitoring without draining battery life remain ongoing concerns.

To address these challenges, this work leverages the power of CNN and LSTM models. CNNs are particularly effective at extracting spatial features from sensor data, while LSTMs are suitable for handling sequential data and capturing temporal dependencies in activity patterns. Together, these models can more accurately recognize complex activities, predict future behaviors, and detect anomalies in real-time.

### 1.1 Contribution of the Research

Motivated by the challenges in existing Human Activity Recognition (HAR) methods, this research makes the following key contributions:

1. The study introduces a comprehensive data preprocessing pipeline, addressing noise,

missing values, and normalization challenges in the WISDM dataset. This ensures the input data is of high quality and suitable for accurate feature learning.

2. A hybrid DL approach is proposed, combining CNN and LSTM for modeling temporal dependencies, effectively capturing complex activity patterns from time-series data.
3. The model demonstrates its effectiveness in classifying diverse human activities, utilizing metrics like precision, recall, and F1-score to validate its effectiveness.

### **1.2 Structure of the Paper**

The sections are organized in the following manner: Section 2 discusses relevant studies conducted by various authors. The dataset preparation, pre-processing, proposed methodology with the detailed description is provided in Section-3. Evaluation outcomes of the suggested approach and comparative assessments are demonstrated in Section-4. In conclusion, the paper discusses future enhancements in Section 5.

## **2. Related Works**

Alarfaj et al. (2024) [13] introduced an innovative technique to human activity recognition utilizing specialized CNN developed for individual sensor types, including accelerometers, gyroscopes, and barometers. Their key contribution was developing sensor-specific CNN models capable of capturing unique characteristics of each sensor type, combined with a late-fusion method to combine predictions from different methods. When evaluated against existing traditional support vector machine (SVM) classifiers, their approach demonstrated superior performance, achieving better accuracy, significantly outperforming the SVM's performance. The research highlighted the importance of integrating multiple sensors with barometer data and using additional filtering algorithms for improved accuracy, though the complexity of managing multiple sensor-specific models may limit real-world implementation.

Alanazi et al. (2024) [14] conducted a comprehensive evaluation of machine learning algorithms for HAR utilising smartphone sensor data. Their study uniquely aimed at testing various algorithms under predefined settings using the KU-HAR dataset. The LightGBM classifier emerged as the superior performer across multiple metrics like accuracy, F1 score, precision, and recall. While the gradient boosting approach demonstrated excellent performance, the study noted its lengthy training time as a significant limitation, though other classifiers achieved acceptable training periods.

Raj and Kos (2023) [15] presented a sophisticated implementation of CNN for HAR tasks using the WISDM dataset. Their approach utilized a two-dimensional CNN model to capture both temporal and spatial data related to human activities. The proposed model achieved an impressive accuracy rate surpassing other ML techniques. Their study highlighted the possibilities of deep learning methods in expanding HAR applications, although it is limited to a single public dataset.

Nia et al. (2023) [16] demonstrated the effectiveness of combining Artificial Neural Networks (ANNs) for feature extraction with the collective decision-making power of Random Forest, resulting in enhanced accuracy for classifying human activities. By utilizing the advantages of both ANN and Random Forest (RF), the integrated model tackles the issues with real, imperfect data, resulting in a more resilient and precise classification system. The outcomes emphasize the efficiency of feature extraction through ANNs and stress the significance of including RF in HAR methods to achieve enhanced accuracy. While the results were promising, the computational complexity of maintaining both ANN and Random Forest components may present challenges in resource-constrained environments.

Li and Wang (2022) [17] focused on HAR for elderly populations, employing an integration

of conventional ML and DL algorithms. Their study utilized smartphone-based gyroscope and accelerometer data to monitor activities in various environments. LSTM achieved the highest accuracy at ,while Support Vector Machine demonstrated efficient performance with 85.07% accuracy and significantly lower computational time of 0.42 minutes utilising 10-fold cross-validation. The research particularly emphasized the importance of efficient activity recognition for elderly care applications. A drawback of this study is the reliance on smartphone sensors, which may not capture all relevant activity patterns or environmental factors.

Nazari et al. (2022) [18] proposed several models for predicting four activities using inertial sensors located in the ankle area of lower-leg assistive device users. Their innovative approach focused on sensors that could be integrated directly into the control unit of assistive devices rather than requiring skin attachment. The Artificial Neural Network-based models achieved average classification accuracy, though the limited scope of activities and sensor placement may restrict broader applications.

Uddin and Soyulu (2021) [19] proposed an innovative time-sequential information-based deep Neural Structured Learning (NSL) system for HAR using body sensor data. Their approach incorporated kernel-based discriminant analysis (KDA) for feature optimization and LSTM for activity modeling, achieving enhanced recall rate on public datasets. Additionally, the system's reliance on specific body sensor data may restrict its applicability across different environments or diverse user populations.

Wang et al. (2020) [20] introduced a wearable sensors-based IOT system for continuous health monitoring of sports persons. Their research implemented various classifiers, with the Ensemble Bayesian deep classifier (EBDC) achieving the best performance with a minimum error rate of 0.25 and maximum recall value, outperforming other approaches including neural networks , RNN, BDNN, and LSTM. Moreover, these models may require substantial computational resources, reducing its scalability for large-scale deployments.

Mahmud et al. (2020) [21] introduced a multi-stage training technique for HAR that uniquely employed multiple transformations on time series data to obtain diverse feature representations. Their method utilized an efficient deep CNN architecture trained on different transformed spaces, achieving remarkable results across multiple datasets: highest accuracy on UCI HAR, USC HAR, and on SKODA database. While the results were impressive, the computational overhead of managing multiple transformation stages and training processes could impact real-time applications.

Hou (2020) [22] designed a comparative assessment of DL and existing ML models for HAR, specifically examining their performance on datasets of varying sizes. The research revealed that traditional machine learning methods performed better on smaller datasets (achieving 87% accuracy), while DL approaches including CNN and LSTM were more effective with larger datasets (reaching 88% accuracy). This research offered valuable findings into the selection of appropriate methodologies based on dataset characteristics, nevertheless its findings may be limited to the specific datasets and architectures tested.

### 3. PROPOSED METHODOLOGY

The proposed HAR system which is shown in Figure 1, is designed as a hybrid deep learning framework comprising four integral components: data collection, data preprocessing, feature extraction, and activity classification. The system utilizes the Wireless Sensor Data Mining (WISDM) dataset, which includes accelerometer and gyroscope data recorded during various human activities. The raw data undergoes rigorous preprocessing to ensure consistency and quality, addressing issues like noise, missing values, and normalization to prepare it for model training. For feature extraction, the system employs a hybrid architecture that combines CNN and LSTM

networks. The CNNs extract intricate spatial features from the time-series data, while the LSTMs capture the temporal dependencies critical for recognizing activity sequences. These features are then transferred to fully connected dense layers for activity classification. This classification component leverages the information-rich features learned by the hybrid model to deliver precise predictions.

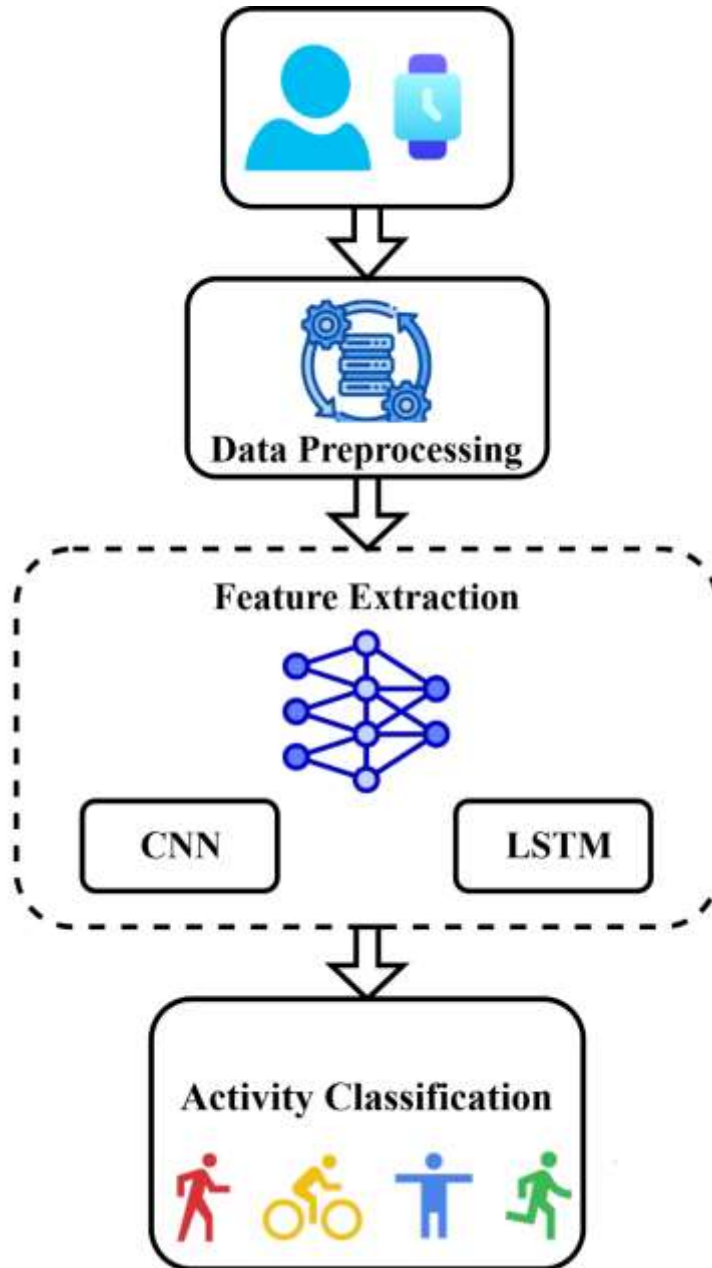


Figure 1: Proposed Architecture

### 3.1 Materials and Methods

For this research, we utilized the WISDM (Wireless Sensor Data Mining) dataset [23], a high-quality time-series dataset specifically designed for human activity recognition (HAR) tasks. It consists of motion data collected from wearable devices using accelerometers during six distinct physical activities: walking, jogging, sitting, standing, upstairs, and downstairs. The dataset contains 1,098,207 records, with class distribution reflecting the natural duration and frequency of each activity. Activities are represented in percentages shown in Figure 2. Data collection was performed at a sampling rate of 20Hz (one sample every 50 milliseconds), capturing fine-grained details of user movements. The dataset is structured with six key attributes, covering user information, activity labels, timestamps, and acceleration data along three axes which are illustrated in table 1. Notably, the acceleration values include gravitational components, making the dataset comprehensive for studying motion patterns.

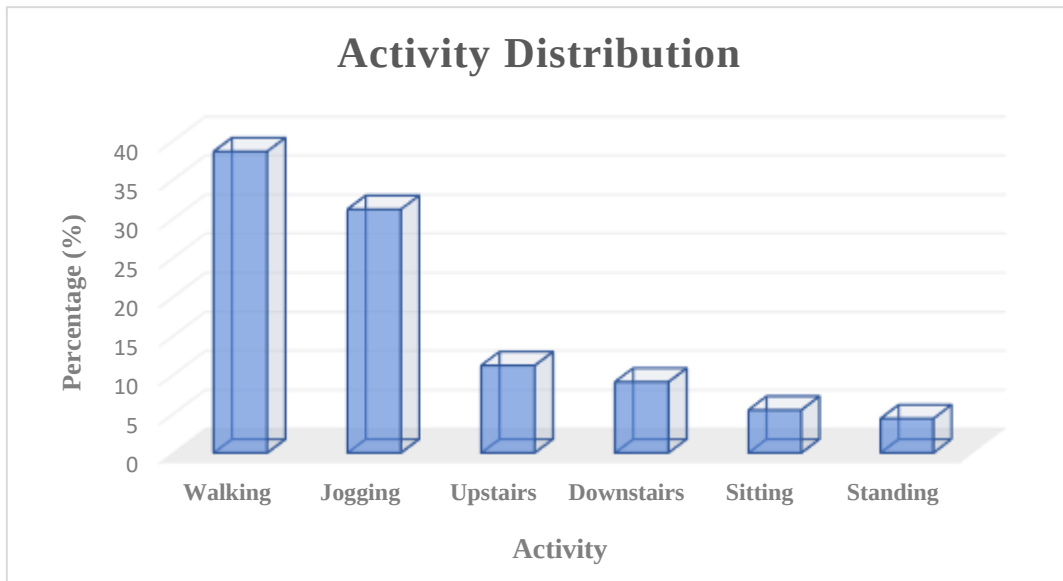


Figure 2: Activity Distribution in the WISDM Dataset

Table 1: Dataset Attribute Description

| Attribute      | Type    | Description                                                                                                        |
|----------------|---------|--------------------------------------------------------------------------------------------------------------------|
| User           | Nominal | Identifier for participants, ranging from 1 to 36.                                                                 |
| Activity       | Nominal | The activity performed, classified into six categories: Walking, Jogging, Sitting, Standing, Upstairs, Downstairs. |
| Timestamp      | Numeric | Device uptime in nanoseconds, representing the timing of recorded motion.                                          |
| x-Acceleration | Numeric | Acceleration along the x-axis in $m/s^2$ , including gravitational acceleration.                                   |
| y-Acceleration | Numeric | Acceleration along the y-axis in $m/s^2$ , including gravitational acceleration.                                   |
| z-Acceleration | Numeric | Acceleration along the z-axis in $m/s^2$ , including gravitational acceleration.                                   |

### **3.2 Data Preprocessing**

The WISDM dataset undergoes key preprocessing steps to ensure its readiness for model training, with a focus on handling missing values and normalizing the data. First, any missing values in the dataset are addressed by removing incomplete records to maintain the integrity of the data. Since missing values can introduce bias or distort model predictions, eliminating these records ensures that only complete and reliable data is used for training. This step guarantees consistency and prevents any gaps that could undermine the effectiveness of the ML techniques.

Next, the dataset is normalized to assure that all input features, such as accelerometer values along the x, y, and z axes, are on a similar scale. Normalization is essential because accelerometer data can vary significantly across different users and activities, leading to discrepancies that could impact the algorithm's proficiency in learning meaningful patterns. In this case, Z-score normalization is applied, which standardizes the data by subtracting the mean and dividing by the standard deviation for each feature. This ensures that the input features have a mean of 0 and a standard deviation of 1, enabling the model to train more efficiently and effectively by preventing any single feature from dominating the learning process.

### **3.3 Feature Extraction Module**

The Feature Extraction Module in this research integrates the benefits of CNN and LSTM networks to examine the raw sensor information collected from wearable devices. CNNs are used to extract intricate spatial features from the accelerometer and gyroscope time-series data, ensuring the identification of activity-specific patterns. LSTMs complement this by capturing the temporal dependencies inherent in the sequential data, enabling the model to understand the progression and context of activities over time. This integrated approach ensures robust and comprehensive feature extraction, which is crucial for achieving high accuracy and reliability in human activity recognition (HAR) tasks.

#### **3.3.1 Convolutional Neural Network (CNN)- An Overview**

It is an DL approach designed to extract features from structured data, like images and time-series data. They are especially powerful for tasks that require localized pattern recognition because of their capacity to self-learn spatial relationships of features. CNN architecture comprises of convolutional layers, pooling layers, and fully connected layers, each contributing to the overall feature extraction process.

##### **3.3.1.1 Convolutional Layers:**

They are the primary components of a CNN. Their role is to apply convolution operations to the input data using filters (also known as kernels). These filters slide over the data, performing a mathematical operation that generates feature maps. These feature maps highlight key characteristics, such as edges or textures, in the input data. In time-series data, convolutional layers capture temporal dependencies and activity-specific features by moving along one dimension. By stacking multiple convolutional layers, a CNN is able to learn progressively more abstract and complex features, enabling the model to represent higher-level patterns and structures from the raw input.

### 3.3.1.2 Pooling Layers:

Pooling layers are placed between convolutional layers to decrease the spatial dimensions of the feature maps. This reduction in size helps minimize computational complexity and prevents the model from becoming overly complex, which could lead to overfitting. The most common type of pooling used in CNNs is max-pooling, where the maximum value from each region of the feature map is chosen. This operation retains the most important features while discarding less relevant data. Pooling not only reduces computational load but also aids in generalizing the model by making it more invariant to slight changes in the input. As a result, pooling layers help to efficiently extract robust and high-level features essential for tasks like activity recognition or object detection.

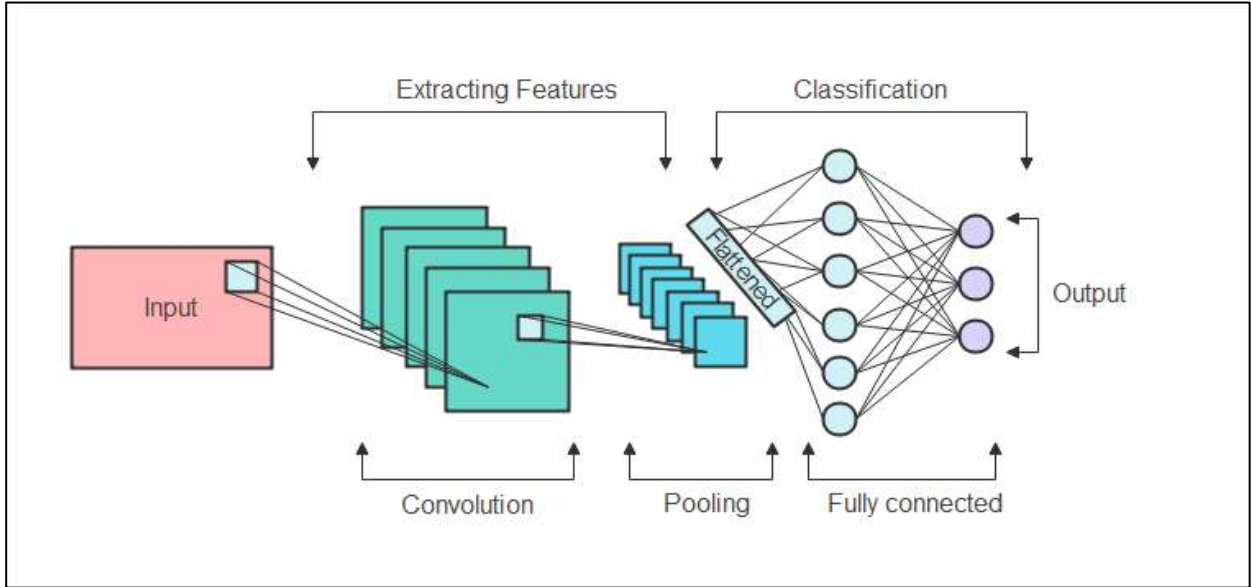


Figure 3: CNN Architecture

The raw time-series data from wearable devices, is first processed through a series of convolutional layers. These layers capture intricate spatial patterns related to human activities, such as walking, running, or sitting. These feature are given to the LSTM layers, which are adept at capturing temporal dependencies.

### 3.3.2 Long Short Term Memory

It is a type of RNN, are integral to the hybrid DL framework proposed for HAR in this system. LSTMs are highly efficient in capturing long-range temporal dependencies, which are essential in accurately recognizing activity sequences in time-series data from accelerometers and gyroscopes. By addressing the challenges of existing RNNs, especially the issue of vanishing gradients, LSTMs facilitate the model in keeping essential information over time while disregarding non-relevant data.

The LSTM architecture as illustrated in Figure 4, is composed of memory cells and gating mechanisms, that control the transmission of information within the network. These components allow LSTMs to effectively handle the sequential nature of human activity data, ensuring accurate recognition across varying time frames.

The fundamental parts of an LSTM unit are:

1. **Cell State ( $C_t$ ):** This represents the memory of the network and carries important

information over time.

2. **Hidden State ( $h_t$ ):** This represents the output of the LSTM unit at time step  $t$ , that is passed to the next layer or time step.
3. **Gates:** These are mechanisms that control the transmission of data within the memory cell. They are:
  - **Forget Gate:** Determines which information to remove from the cell state.
  - **Input Gate:** Determines which new information to save in the cell state.
  - **Output Gate:** Decides what the next hidden state will be.

#### 3.3.2.1. Forget Gate ( $f_t$ )

It controls how much of the previous cell state ( $C_{t-1}$ ) should be carried forward to the current cell state. It takes the previous hidden state ( $h_{t-1}$ ) and the current input ( $x_t$ ) as inputs and outputs a value between 0 and 1, which is then used to scale the previous cell state.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

Here  $f_t$  is the forget gate's output,  $W_f$  is the weight matrix for the forget gate,  $b_f$  is the bias term and  $\sigma$  is the sigmoid activation function, which squashes the output between 0 and 1.

#### 3.3.2.2. Input Gate ( $i_t$ ):

It decides how much of the new information should be incorporated into the cell state. The candidate cell state signifies the possible new data that can be included in the cell state. The input gate is measured as follows:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

Here  $i_t$  is the input gate's output,  $W_i$  is the weight matrices for the input gate and  $b_i$  is the bias terms.

#### 3.3.2.3. Update Cell State ( $C_t$ ):

This is modified by merging the previous cell state ( $C_{t-1}$ ) and the new candidate cell state ( $\widetilde{C}_t$ ), managed by the forget gate and the input gate. The update rule is as follows:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t \quad (3)$$

Where  $C_t$  is the updated cell state,  $f_t$  regulates the extent to which the previous cell state is preserved and  $i_t$  manages the degree to which the new candidate cell state is incorporated into the cell state.

#### 3.3.2.4. Output Gate ( $o_t$ ) and Hidden State ( $h_t$ ):

The output gate decides what the next hidden state ( $h_t$ ) will be. The hidden state is based on the updated cell state ( $C_t$ ), which is passed through the output gate and squashed using the tanh function. The output gate is computed as follows:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (5)$$

Where  $o_t$  is the output gate's output,  $h_t$  is the hidden state,  $W_o$  is the weight matrix for the output gate,  $b_o$  is the bias term and the tanh ( $C_t$ ) function squashes the cell state to ensure that the hidden state is between -1 and 1.

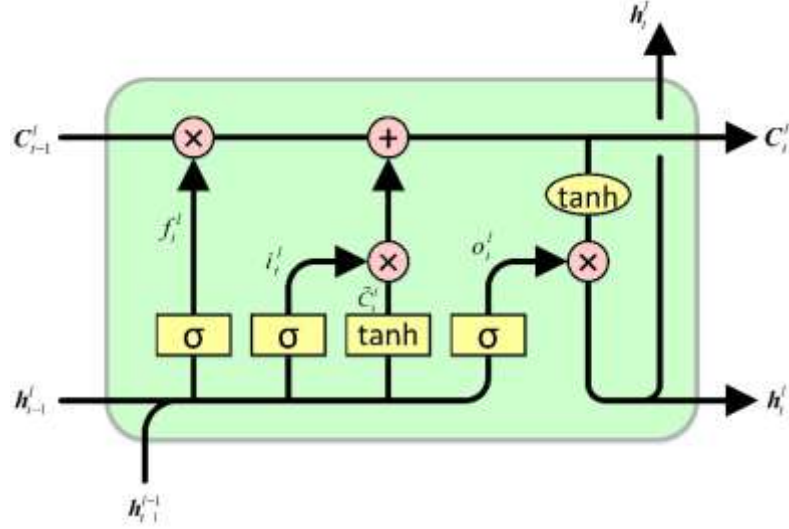


Figure 4: LSTM Architecture

These components enable the LSTM to capture long-term dependencies in time-series data by maintaining and updating relevant information across time steps. This ability makes it particularly well-suited for recognizing complex and evolving activity patterns, essential for accurate human activity recognition.

### 3.4 Dense Classification Layers

The dense classification layers in the proposed human activity recognition (HAR) system serve as the final decision-making stage, where the features extracted by the CNN and LSTM layers are transformed into predictions for different activity classes. After the hybrid model captures spatial features through CNNs and temporal dependencies through LSTMs, the extracted feature vector is passed to the fully connected layers, or dense layers. These layers execute a linear transformation of the feature vector, followed by an activation function, which helps capture complex relationships between the learned features.

For multiclass classification, each neuron in the final dense layer corresponds to one of the activity classes. The output of these neurons represents the network's confidence in each class, with the highest value indicating the predicted activity. The dense layers are followed by a softmax activation function, which is particularly suitable for multiclass problems as it normalizes the outputs into a probability distribution.

The softmax function is applied to the output of the dense layer to convert the raw outputs (logits) into a probability distribution over all the activity classes. These classification layers, equipped with softmax, ensure that the network provides accurate and interpretable predictions for the multiple activity classes in the WISDM dataset.

## 4. Results and discussion

This section provides the evaluation of the proposed human activity recognition system, highlighting key performance metrics and discussing the performance of the hybrid CNN-LSTM in classifying activities from the WISDM dataset.

### 4.1 Implementation Details

The model was developed utilising Python3.19 programming and libraries like matplotlib, numpy, pandas, Scikit-Learn, seaborn are utilized for evaluating the proposed model. The experimentation was carried out in the PC workstation with i7 CPU, 3.2 GHZ operating frequency, 16GB RAM, NVIDIA Tesla GPU.

#### 4.2 Evaluation Metrics

Evaluation metrics like accuracy, precision, recall, specificity as well as F1-score are measured and immediately evaluated against other state of the art DL models to show the effectiveness of the proposed model. These measures also highlight the model's productivity and low computational overhead. Table 2 displays the mathematical expression employed to calculate the performance metrics. The early halting method is applied to avoid overfitting and generalization problems. This technique is used to terminate the iteration when the validation performance of the recommended model fails to improve over time.

Table 2: Evaluation Metrics utilized for assessment

| SL.NO | Performance Measures | Expression                                              |
|-------|----------------------|---------------------------------------------------------|
| 1     | Accuracy             | $\frac{TP + TN}{TP + TN + FP + FN}$                     |
| 2     | Recall               | $\frac{TP}{TP + FN} \times 100$                         |
| 3     | Specificity          | $\frac{TN}{TN + FP}$                                    |
| 4     | Precision            | $\frac{TP}{TP + FP}$                                    |
| 5     | F1-Score             | $2 \cdot \frac{Precision * Recall}{Precision + Recall}$ |

TP & TN are True Positive & negative, FP & FN are False Positive& negative.

#### 4.3 Experimental Outcomes

Table 3 presents a comparative assessment of various DL algorithms for HAR based on key evaluation metrics such as accuracy, sensitivity, specificity, precision, and F1-score. The proposed hybrid model outperforms the baseline models, including Artificial Neural Networks (ANN), Gated Recurrent Units (GRU), LSTM and CNN. With an accuracy of 99.4%, the proposed algorithm highlights high performance for all metrics, highlighting its robustness in recognizing activities from the WISDM dataset.

Table 3: Comparative Analysis between the Different Models in recognizing activities.

| Algorithm      | Accuracy (%) | Sensitivity (%) | Specificity (%) | Precision (%) | F1-Score (%) |
|----------------|--------------|-----------------|-----------------|---------------|--------------|
| ANN            | 84.6%        | 85.8%           | 83.6%           | 83.4%         | 85.9%        |
| GRU            | 86.1%        | 86.4%           | 85.6%           | 85.8%         | 85.8%        |
| LSTM           | 88.3%        | 87.5%           | 86.7%           | 86.7%         | 87.8%        |
| CNN            | 92%          | 91.3%           | 89.1%           | 90%           | 91.8%        |
| Proposed Model | 99.4%        | 98.6%           | 99.3%           | 99%           | 98.2%        |

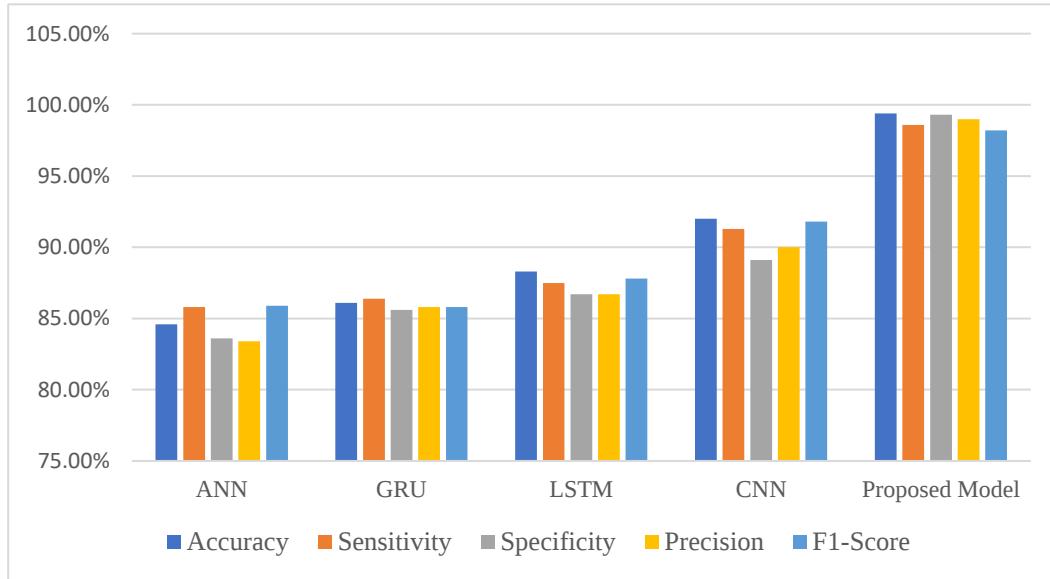


Figure 5: Comparative Analysis with Varying DL models based on testing accuracy

Figure 5 illustrates the comparative assessment of testing accuracy among various deep learning methods. It highlights the consistent superiority of the proposed hybrid CNN-LSTM model over other architectures, with a significant margin in performance accuracy.

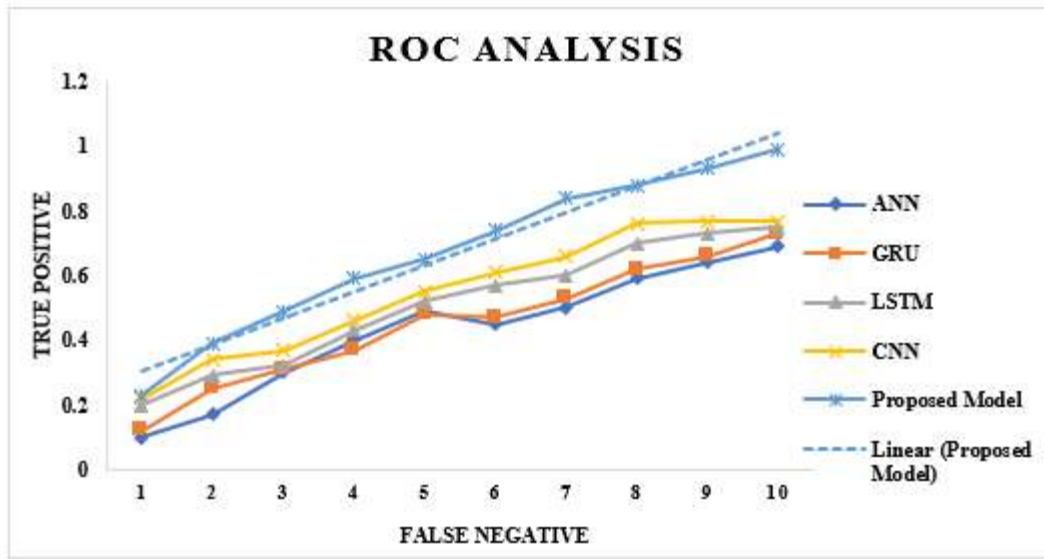


Figure 6: ROC Curves for the Different Models in recognizing activities.

Figure 6 shows the ROC curves for different models, showcasing the ability of each model to differentiate between activity classes. The proposed model achieves an enhanced ROC curve, demonstrating excellent classification performance.

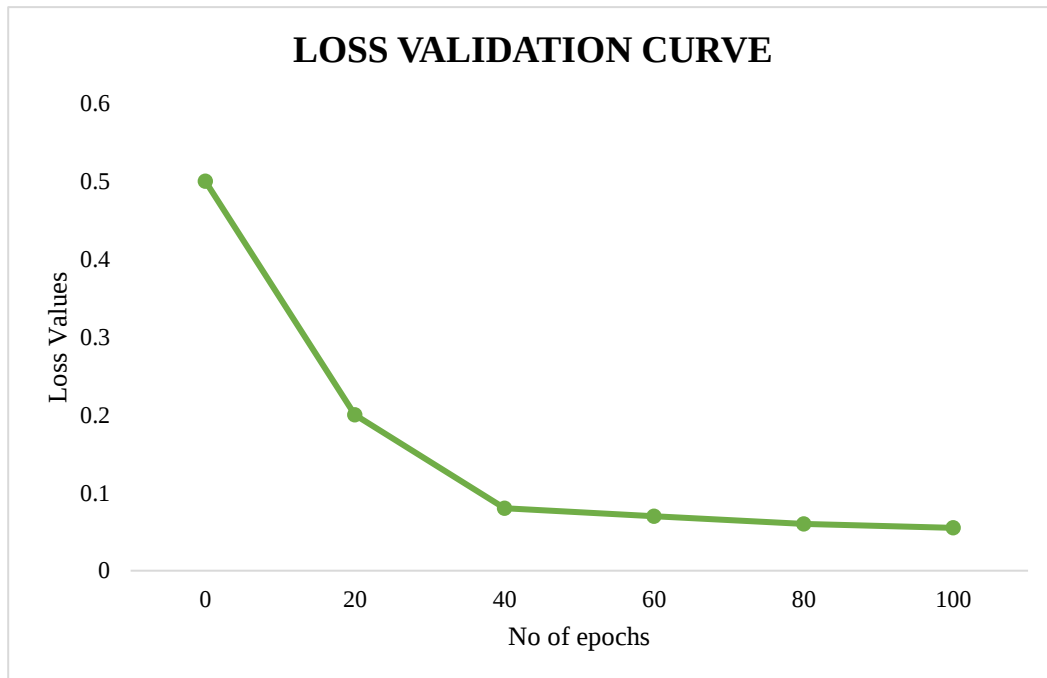


Figure 7: Loss Validation Curve for the Proposed model

Figure 7 depicts the loss validation curve for the recommended approach during training. It highlights the convergence of the model, with the validation loss decreasing steadily and stabilizing, indicating the model's ability to generalize well without overfitting. These results collectively emphasize the performance and reliability of the recommended hybrid model in human activity recognition tasks.

## 5. CONCLUSIONS

This research underscores the effectiveness of integrating wearable IoT devices with advanced AI techniques to design a robust Human Activity Recognition (HAR) system. Wearable IoT devices, equipped with sensors like accelerometers and gyroscopes, continuously collect diverse data on human movements, offering a rich foundation for activity analysis. By utilizing the WISDM dataset, the proposed hybrid deep learning framework—combining CNNs for spatial feature extraction and LSTMs for temporal sequence modeling—achieved an improved classification accuracy alongside strong evaluation metrics such as precision, recall, specificity, and F1-score. The model successfully addresses challenges like data variability and noise, showcasing its reliability in real-time activity analysis. Its robust performance highlights its practical utility in applications such as personalized healthcare, fitness monitoring, and remote health systems. The study not only advances state-of-the-art HAR models but also demonstrates the feasibility of deploying such systems in real-world environments, setting a foundation for further innovation in wearable IoT-based activity analysis. Future research can explore integrating additional sensors, such as physiological and environmental ones, to enhance activity recognition. Incorporating energy-efficient models and hardware optimization will further improve real-time applicability in resource-constrained settings. Expanding datasets to include diverse populations and activities will improve model generalization and inclusivity.

---

**Funding Statement:** The authors received no specific funding for this study.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

---

## References

- [1] A. Khanna and S. Kaur, "Internet of Things (IoT) Applications and Challenges: A Comprehensive Review", *Wirel. Pers. Commun.*, vol. 114, pp. 1-76, Sep. 2020.
- [2] Vikram Rajpoot; Sudeep Ray Gaur; Aditya Patel; Akash Saxena, "Internet of Things for Human-Activity Recognition Based on Wearable Sensor Data," in *Sensor Data Analysis and Management: The Role of Deep Learning*, IEEE, 2021, pp.19-31, doi: 10.1002/9781119682806.ch2.
- [3] Paraschiakos S., Cachucho R., Moed M., van Heemst D., Mooijaart S., Slagboom E. P., Knobbe A., and Beekman M., Activity recognition using wearable sensors for tracking the elderly, *User Modeling and User-Adapted Interaction*. (2020) 30, no. 3, 567–605, <https://doi.org/10.1007/s11257-020-09268-2>.
- [4] Dian, F. & Vahidnia, Reza & Rahmati, Alireza. (2020). Wearables and the Internet of Things (IoT), Applications, Opportunities, and Challenges: A Survey. *IEEE Access*. PP. 1-1. 10.1109/ACCESS.2020.2986329.
- [5] Mukhopadhyay SC, Suryadevara NK, Nag A. Wearable Sensors and Systems in the IoT. *Sensors*. 2021; 21(23):7880. <https://doi.org/10.3390/s21237880>
- [6] Rahmani, Amir & Szu-Han, Wang & Yu-Hsuan, Kang & Haghparast, Majid. (2022). The Internet of Things for Applications in Wearable Technology. *IEEE Access*. PP. 1-1. 10.1109/ACCESS.2022.3224487.
- [7] G. S. Mubibya and J. Almhana, "Improving Human Activity Recognition using ML and Wearable Sensors," *ICC 2022 - IEEE International Conference on Communications*, Seoul, Korea, Republic of, 2022, pp. 165-170, doi: 10.1109/ICC45855.2022.9839267.

- [8] A. Subasi, K. Khateeb, T. Brahimi and A. Sarirete, "Human activity recognition using machine learning methods in a smart healthcare environment" in *Innovation in Health Informatics*, Elsevier, pp. 123-144, 2020.
- [9] D. Radivojević, N. Mirkov and S. Maletić, "Human Activity Recognition based on Machine Learning Classification of Smartwatch Accelerometer Dataset", *FME Trans.*, vol. 49, pp. 225-232, Jan. 2021.
- [10] Sabry, F., Eltaras, T., Labda, W., Alzoubi, K., & Malluhi, Q. (2022). Machine Learning for Healthcare Wearable Devices: The Big Picture. *Journal of healthcare engineering*, 2022, 4653923. <https://doi.org/10.1155/2022/4653923>
- [11] Mekruksavanich S. and Jitpattanakul A., LSTM networks using smartphone data for sensor-based human activity recognition in smart homes, *Sensors*. (2021) 21, no. 5.
- [12] Xinyu Song, Hongyu Sun, Yanhua Dong, Ying Pei, and Shuai Wang. 2024. Deep Learning-Based Wearable Human Activity Recognition: Model and Performance Analysis. In *Proceedings of the 2024 8th International Conference on Control Engineering and Artificial Intelligence (CCEAI '24)*. Association for Computing Machinery, New York, NY, USA, 30–36. <https://doi.org/10.1145/3640824.3640830>
- [13] Alarfaj, M., Al Madini, A., Alsafran, A., Farag, M., Chtourou, S., Afifi, A., Ahmad, A., Al Rubayyi, O., Al Harbi, A., & Al Thunaian, M. (2024). Wearable sensors based on artificial intelligence models for human activity recognition. *Frontiers in Artificial Intelligence*, 7. <https://doi.org/10.3389/frai.2024.1424190>.
- [14] Alanazi, M., Aldahr, R.S. and Ilyas, M. 2024. Human Activity Recognition through Smartphone Inertial Sensors with ML Approach. *Engineering, Technology & Applied Science Research*. 14, 1 (Feb. 2024), 12780–12787. DOI: <https://doi.org/10.48084/etasr.6586>.
- [15] Raj, R., Kos, A. An improved human activity recognition technique based on convolutional neural network. *Sci Rep* 13, 22581 (2023). <https://doi.org/10.1038/s41598-023-49739-1>.
- [16] Nia, Nafiseh & Amiri, Amin & Nasab, Ahad & Kaplanoglu, Erkan & Liang, Yu. (2023). The Power of ANN-Random Forest Algorithm in Human Activities Recognition Using IMU Data. 1-7. 10.1109/BHI58575.2023.10313507.
- [17] Li, Y., & Wang, L. (2022). Human activity recognition based on residual network and BiLSTM. *Sensors*, 22(2), 635.
- [18] F. Nazari, D. Nahavandi, N. Mohajer and A. Khosravi, "Comparison of Deep Learning Techniques on Human Activity Recognition using Ankle Inertial Signals," 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Prague, Czech Republic, 2022, pp. 2251-2256, doi: 10.1109/SMC53654.2022.9945472.
- [19] Uddin, M. Z., & Soylu, A. (2021). Human activity recognition using wearable sensors, discriminant analysis, and long short-term memory-based neural structured learning. *Scientific reports*, 11(1), 16455. <https://doi.org/10.1038/s41598-021-95947-y>.
- [20] Wang, H., Kadry, S. N., & Raj, E. D. (2020). Continuous health monitoring of sportsperson using IoT devices based wearable technology. *Computer Communications*, 160, 588–595. <https://doi.org/10.1016/j.comcom.2020.04.025>.
- [21] Mahmud, T., Sayyed, A. S., Fattah, S. A., & Kung, S. Y. (2020). A novel multi-stage training approach for human activity recognition from multimodal wearable sensor data using deep neural network. *IEEE Sensors Journal*, 21(2), 1715-1726.
- [22] Hou, C. (2020). A study on IMU-Based Human Activity Recognition Using Deep Learning and Traditional Machine Learning. 2020 5th International Conference on Computer and Communication Systems (ICCCS), 225-234.
- [23] Seelwal, P., Prasad, A. N., Srinivas, C., & Rekha, V. S. (2023). Human activity recognition using WISDM datasets. *Journal of Online Engineering Education*, 14(1s), 88.