Received: September 2, 2023. Revised: October 10, 2023.

INASS

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

A Novel Framework for Future Human Activity Prediction Using Sensor-Based Data

Mohammad Khalaf Rahim Al-juaifari[1]* **Alih. Ali Athari[2]**

[1] *Department of computer Science College of Computer Science/ and math/ University of Kufa, Najaf, Iraq*
[2] *Department of Communication and Electronic, college of Engineer, University of Kufa, Najaf, Iraq*
* Corresponding author's Email: mohammad.aljuaifari@uokufa.edu.iq

Abstract: The prediction of human activities has garnered significant attention, owing to their relevance in diverse applications spanning healthcare, robotics, and user-computer interaction. This paper addresses the imperative need for a comprehensive multi-step prediction system tailored to accommodate these varied domains. The proposed framework comprises three key stages. Firstly, wavelet transform (WT) is employed for data pre-processing to eliminate noise and render the data amenable for time series analysis, the second phase of the framework involves the utilization of a hybrid model combining Long Short-Term Memory (LSTM) and Convolutional 1D (CONV1D) layers, denoted as LSTM-CONV1D designed to effectively address the complexities involved in extracting relevant features and tackling data imbalance issues. The LSTM component is employed for human activity classification based on sensor data, taking into account the sequential nature of activities. Concurrently, the CONV1D component is utilized for feature extraction. In the third phase, a method is introduced for predicting future activity levels and steps. This method incorporates a function that takes the output of the trained LSTM-CONV1D model as an input sequence. The performance of the proposed model is rigorously evaluated across four benchmark datasets based on sensor data: UCI-HAR (Accuracy: 99.03%), M-health (Accuracy: 93.21%), WISDM (Accuracy: 89.43%), and PAMAP2 (Accuracy: 86.85%). Notably, these results represent the state-of-the-art performance characterized by its simplicity attributed to the utilization of a single-layer LSTM and CONV1D. Furthermore, an additional evaluation is conducted on the UCI-HAR dataset with an accuracy of 98.92, wherein the LSTM time series function iteratively generates predictions for future time steps, following the same pre-processing steps.

Keywords: Deep learning (DL), Human activity recognition (HAR); Convolution neural network (CNN); Long short-term memory (LSTM); Machine learning (ML), Wavelet transform (WT), Human activity prediction (HAP).

1. Introduction

Predicting future activity in the absence of extensive data poses a formidable challenge, particularly in domains like human movement prediction. Existing classification models predominantly focus on forecasting near-term outcomes while largely neglecting the explicit consideration of future events. To address this crucial gap, the development of a scalable model capable of adapting to temporal and spatial variations and accounting for the correlation among sequential human movements is imperative. The UCI-HAR dataset, a widely recognized benchmark in this domain, serves as a pivotal foundation for our research benchmarks. (WT) facilitates the examination of signals at varying resolutions by decomposing time-series data into diverse scales and frequencies. Within the context of UCI-HAR, (WT) plays a pivotal role in extracting activity-specific features and enabling the classification of activities undertaken by subjects. Predicting the future, especially when confronted with multiple probable scenarios, presents another formidable challenge.

Consequently, modeling sequential dependencies within data while concurrently predicting various conceivable long-term predictions becomes a pressing concern, several factors encompass data

Received: September 2, 2023. Revised: October 10, 2023.

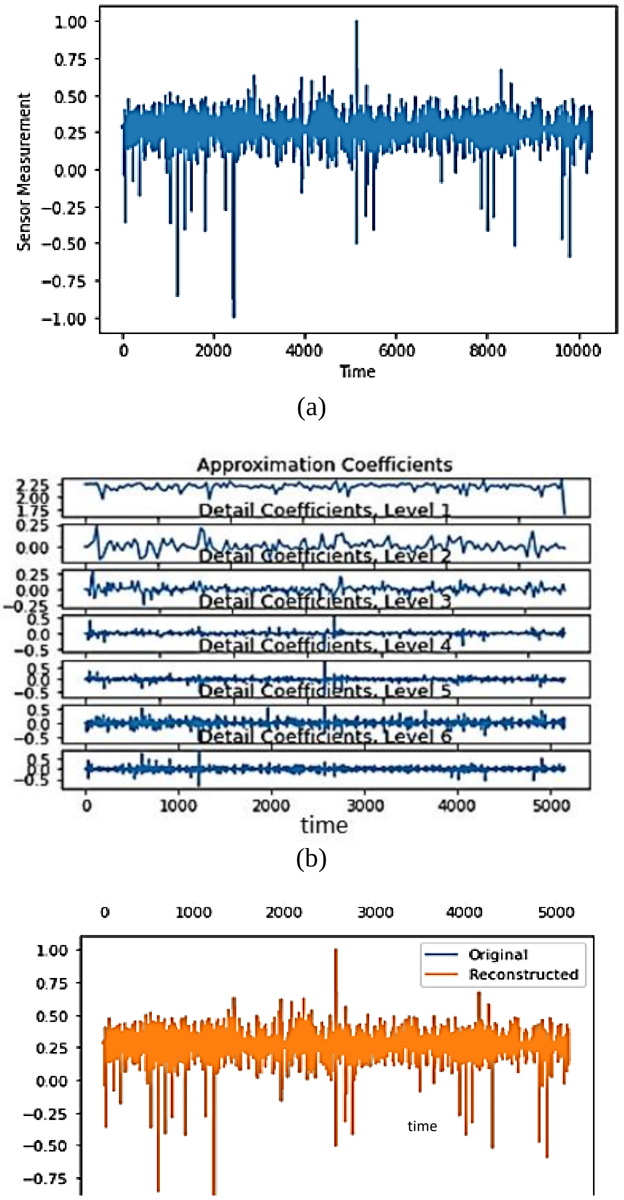


Figure. 1: (a) UCI-HAR analysis sensor measurement "original signal" and (b) Approximation and detailed coefficients, and (c) Reconstructed signal using (WT)

complexity, dataset size, computational resources, and performance criteria. This paper introduces several contributions to the field. Firstly, it pioneers the application of a potent feature extraction method previously unexplored in the realm of human activity prediction. Secondly, we present the architecture of a robust model that combines (LSTM) and (CONV1D) models trained on four distinct datasets. Lastly, we propose a novel framework that incorporates a temporal-spatial model featuring an (LSTM) layer, specifically designed for future event prediction. In our approach, (DWT) is applied to the original sensor data, as depicted in Fig. 1. This process involves transforming data generated by sensors (a) into approximation coefficients with detailed levels (b) and subsequently reconstructing the signal using (DWT) (c). This procedure eliminates noise while preserving the temporal sequence inherent in time series data.

The key contributions include:

1-Unified multidata set framework: It can predict human activities across four different datasets. It employs consistent methods for preprocessing and modeling, enhancing versatility and facilitating fair comparisons.

2- Time complexity reduction: Single layer model optimized for reducing time complexity. This model enables quick identification of the next step in human activities, enhancing real-world applicability.

3- Versatile and efficient framework: The proposed framework comprises two sections—one for predicting future activities using the UCI-HAR.

2. Literature review

Predicting future events presents challenges, particularly when data is limited. Consequently, classification models often focus on short-term predictions, such as human movements, without explicitly addressing future events. To bridge this gap, the need arises for a scalable model that can adapt to temporal and spatial variations and effectively capture the sequential correlation of human movements.

In their research, the authors [1] offer a definition for gait analysis, defining it as the study of human locomotion. This analysis is facilitated by accelerometers, which quantify linear acceleration related to physical activity, and gyroscopes, which measure angular velocity [2]. The authors employ a hybrid model combining (CNN) and Bidirectional (BiLSTM) for classification, achieving an impressive accuracy of 96%. (HAR) [3] is depicted as a pattern recognition task, encompassing four crucial phases: data collection, data pre-processing, feature extraction, and activity classification. Another study by authors [4] introduces a CNN-LSTM hybrid model to enhance smartphone-based human activity recognition accuracy. Fine-tuning hyperparameters through Bayesian optimization results in a 2.24% improvement in classification accuracy. In [5], an ensemble learning algorithm utilizing smartphone sensor data achieved a prediction accuracy of 96.7% using the F1 measure. Additionally, [6] proposes a Two-Dimensional (2D-CNN) for qualitative human activity recognition, achieving accuracy ranging from 96.57% to 99.28% along with high Cohen's kappa values.

 DOI: 10.22266/ijies2023.1231.81

Authors in [7] present a model involving a deep neural network-based CNN and a gated recurrent unit (CNN-GRU) and evaluate its performance using the UCI-HAR, WISDM, and PAMAP2 datasets, yielding accuracy rates of 96.20%, 97.21%, and 95.27%, respectively. Parallel LSTM layers combined with the convolution layer, as proposed in [8], demonstrate their effectiveness. Moreover, [9] introduces a CNN-based approach for human activity recognition with the WISDM dataset, achieving a prediction accuracy of 89.67% across six distinct activities. In [10], an optimization technique cantered on crop prediction surpasses conventional approaches when tested with public datasets. The rise of wearable devices like smartwatches and wristbands has enabled the capture and analysis of individuals' daily physical activity, transcending environmental constraints [11]. Furthermore, the integration of fusion techniques into classification models [12] enhances accuracy using cardiac datasets. However, traditional approaches encounter challenges in representing intricate activities and processing multi-channel data due to their limited adaptability. Consequently, (DL) techniques, including (CNNs) [14, 15], have garnered increased attention in this domain.

3. Methodology and system design:

Methods and sequence of procedures to be proposed for doing our framework are described in this section:

3.1 Sensor-based datasets:

The following datasets are used to train our models for predicting the next step of activity based on sensor data from a wearable device:

A -UCI human activity recognition using smartphones dataset: This dataset contains recordings of 30 subjects performing six different activities (walking, walking upstairs, walking downstairs, sitting, standing, and laying) while wearing a smartphone on their waist. The dataset includes raw sensor data from the phone's accelerometer and gyroscope and pre-processed features such as mean, standard deviation, and correlation coefficients.

B-WISDMV2.0 (wireless sensor data mining) dataset: This dataset contains accelerometer and gyroscope data from both smartphones and smartwatches, collected from 51 participants performing six different activities (walking, jogging, climbing stairs, sitting, standing, and lying down). The dataset includes raw sensor data as well as pre-processed features such as mean, standard deviation, and frequency-domain features.

Table 1 Comparison of model with datasets

Reference	Proposed Model	Dataset (s)	Performance Metrics	Hyperparameters
xxxxx2023 [16]	An automated CNN	UCI-HAR	average accuracy of 98.5 (∓1.1)	nb_filters nb_blocks Alpha
2022 [17]	CNN-LSTM, Self-Attention	MHEALTH and UCI-HAR	Accuracy 98.76%93.11	two attention, batch normalization, dropouts, and dense
2023 [18]	Bi-LSTM	Mhealth, PAMAP2	95.79 and 93.41	No
[19] 2021	CNN + Bi-LSTM	WISDM, UCI - HAR	Accuracy = 98.53, 98	No

C-Mhealth (mobile health): dataset is a collection of sensor data collected from a wearable sensor system that was used to monitor the physical activity of patients suffering from Parkinson's disease. The dataset includes accelerometer and gyroscope data, as well as heart rate data, collected from 10 patients over a period of several months.

D-PAMAP2 (physical activity monitoring dataset): This dataset contains accelerometer, gyroscope, and magnetometer data collected from a wearable sensor system worn by 9 subjects performing 18 different activities, including basic daily activities (e.g., sitting, standing, walking) and sports activities (e.g., cycling, rowing, boxing).

Table 1 shows some use cases, with one or two dataset out of all results and methods used.

3.2 Apply window function to (DWT):

The pre-processing procedure applies the window function to of (DWT) based four datasets that apply equations. The following detailed steps applying:

```
|----> Load the dataset

|        |----> Drop columns of  'activity' and 'Participants'

|     |----> Scale the data using Min-Max scaling
```

```
|       |----> Apply discrete wavelet transform

|       |       |----> Apply a window function (Hann, Hamming, or Blackman)

|     |      |----> Perform the wavelet transform using sym4 wavelet

|       |----> Split the data into training and test sets

|       |----> Encode the labels using LabelEncoder

|         |----> Reshape the input data to include an additional dimension

|----> Build the model

|         |----> Add LSTM layer with 128 units and 'relu' activation

|          |----> Add Conv1D layer with 128 filters, kernel size=5, and 'relu' activation

|          |----> Add MaxPooling1D layer with pool size=2

|        |----> Add Flatten layer to convert 2D data to 1D

|      |----> Add Dense layer with 128 units and 'relu' activation

|       |----> Add BatchNormalization layer

|       |----> Add Dropout layer with dropout rate of 0.3

|      |----> Add Dense layer with softmax activation for multi-class classification

|        |compile, learning rate (lr) declaration then traing and adjust (lr).

|----> Predict on test set and convert one-hot encoded predictions to class labels
```

3.2.1. Windowing and (DWT) for UCI-HAR:

Eq. (1) shows the mathematical of the proposed Pre-processing as follows:

$$DWT(x, \psi j,k, \varphi j,k) = \Sigma k\ \Psi j,k * (x, \psi j,k) + \Sigma k\ \Phi j,k * (x, \varphi j,k) \qquad (1)$$

In this equation, x represents the input signal, which can be a time-domain signal from the UCI-HAR dataset, $\psi j,k$ represents the wavelet function (e.g., Daubechies wavelet) at scale j and translation k. This wavelet function captures the detail coefficients, representing high-frequency components of the signal. $\varphi j,k$ represents the scaling function (low-pass filter) at scale j and translation k. This scaling function captures the approximation coefficients, representing the low-frequency components of the signal. $(x, \psi j,k)$ and $(x, \varphi j,k)$ denotes the inner product of the input signal x with the wavelet and scaling functions, respectively. The equation computes the DWT by decomposing the input signal into a linear combination of wavelet and scaling functions at different scales and translations. The resulting coefficients represent the signal's frequency content at different resolutions. DWT is often implemented using efficient algorithms such as the Mallat algorithm or the lifting scheme. To adapt the equation to the UCI-HAR dataset, replace x with the specific time-domain signal from the dataset and select appropriate wavelet $(\psi j,k)$ and scaling $(\varphi j,k)$ functions based on analysis requirements.

3.2.2. (DWT) for m-health dataset:

Eq. (2) shows the mathematical of the proposed pre-processing is as follows:

$$DWT(x, \psi j,k, \varphi j,k) = \Sigma k\ \psi j,k * (x, \psi j,k) + \Sigma k\ \varphi j,k * (x, \varphi j,k) \qquad (2)$$

In this equation: The input signal, denoted by x, represents the time-domain data from the m-health dataset. The wavelet function at scale j and translation k is represented by $\psi j,k$. This function captures the detail coefficients, indicating high-frequency components. The scaling function at scale j and translation k is represented by $\varphi j,k$. This function captures the approximation coefficients, reflecting low-frequency components.

$(x, \psi j,k)$ and $(x, \varphi j,k)$ represents the inner product of the input signal x with the wavelet and scaling functions, respectively.

By applying this equation, the (DWT) decomposes the input signal into a linear combination of wavelet and scaling functions at different scales and translations. The resulting coefficients provide information about the signal's frequency content at varying resolutions. The Mallat algorithm or the lifting scheme are commonly used for implementing the DWT. To adapt this equation to the m-health dataset, substitute x with the specific time-domain signal from the dataset, and select appropriate wavelet (ψj,k) and scaling (φj,k) functions based on specific analysis objectives.

3.2.3. Apply windowing to (DWT) for WISDM:

Eq. (3) shows the mathematical of the proposed Pre-processing as follows:

$$DWT(x, \psi j,k, \varphi j,k) = \Sigma k\ \psi j,k * (x, \psi j,k) + \Sigma k\ \varphi j,k * (x, \varphi j,k) \quad (3)$$

the input signal, denoted as x, represents the time-domain dataset time series dataset. The wavelet function at scale j and translation k is denoted as ψj,k. This function captures the detail coefficients, which represent high-frequency components of the signal. The scaling function at scale j and translation k is denoted as φj,k. This function captures the approximation coefficients, representing the low-frequency components of the signal. The notation (x, ψj,k) and (x, φj,k) represents the inner product of the input signal x with the wavelet and scaling functions, respectively. By applying this equation, the (DWT) decomposes the input signal into a linear combination of wavelet and scaling functions at different scales and translations. This decomposition provides valuable information about the frequency content of the signal at various resolutions. It is important to note that efficient algorithms like the Mallat algorithm or the lifting scheme are commonly employed for the practical implementation of the DWT. To utilize this equation with the WISDM time series dataset, substitute x with the specific time-domain signal from the dataset, and select appropriate wavelet (ψj,k) and scaling (φj,k) functions based on the specific requirements and goals of the analysis.

3.2.4. PAMAP2 (Physical Activity Monitoring Using Smartphones) Dataset

Eq. (4) shows the mathematical of the proposed Pre-processing as follows:

$$DWT(x, \psi j,k, \varphi j,k) = \Sigma k\ \psi j,k * (x, \psi j,k) + \Sigma k\ \varphi j,k * (x, \varphi j,k) \quad (4)$$

The input signal, denoted as x, represents the time-domain data from the PAMAP2 dataset. The wavelet function at scale j and translation k is represented by ψj,k. This function captures the detail coefficients, which reflect high-frequency components of the signal. The scaling function at scale j and translation k is represented by φj,k. This function captures the approximation coefficients, representing the low-frequency components of the signal. The notation (x, ψj,k) and (x, φj,k) indicates the inner product of the input signal x with the wavelet and scaling functions, respectively. By applying this equation, the DWT decomposes the input signal into a linear combination of wavelet and scaling functions at different scales and translations. This decomposition provides valuable information about the signal's frequency content at various resolutions. The Mallat algorithm or the lifting scheme is commonly used for the practical implementation of the DWT. To adapt this equation to the PAMAP2 dataset, substitute x with the specific time-domain signal from the dataset, and select suitable wavelet (ψj,k) and scaling (φj,k) functions based on the particular analysis goals and requirements.

3.3 Framework design:

Build a system for predicting future human activities using wearable sensors with (DL), there are several components involved, as well as system architecture. The exact implementation details may vary depending on the application and available resources. In general, the structure of the system is described in Fig. 1, LSTM layer to output a sequence rather than a single value, in order to predict the labels of the future sequence that contains the (k+1)th, (k+2)th, and (k+3)th activities, the utilize activity features extracted from the kth, (k-1)th, and (k-2)th activities. Similarly, when forecasting the labels of the future sequence that contains the (k+2)th, (k+3)th, and (k+4)th activities, the use of activity features obtained from the (k+1)th, kth, and (k-1)th activities, the use of activity features obtained from the (k+1)th,(k+2)th, and kth activities and so on. The proposed structure of the system is described in Figure 1: and represents the hidden state hi for each time step. The Conv1D layer conducts a convolution operation between the input and a set of filters, followed by a ReLU activation function. LSTM unit in the layer contains a set of internal hidden state variables, including cell state c(t) and hidden state h(t), which are updated at each time step t based on the input x(t) and the previous state h(t-1).

A refers to activity, i represents the loop for all activities, j is used to group similar activities, for ki collects all related actions as activity groups Activity 1= Aj1+j2+…+j(i-1) +1 to Activity k= Aj1+j2+…+j(i-1) +ni. and for kk collects all related activity as separated groups: Activity 1= Ak1+jk2+…+j(k-1) +1 to Activity k= Ak1+k2+…+j(k-1) +n(k). Finally prediction can be mathematically calculation as: Prediction = Aj1+j2+…+j(i-1)+x

Received: September 2, 2023. Revised: October 10, 2023.

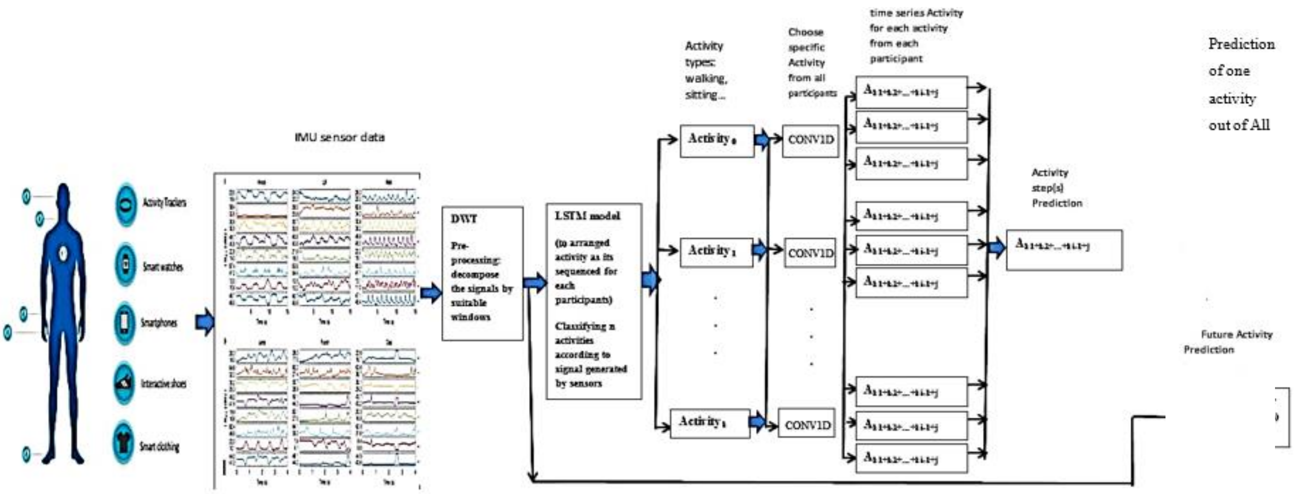


Figure. 2 Architecture of proposed framework

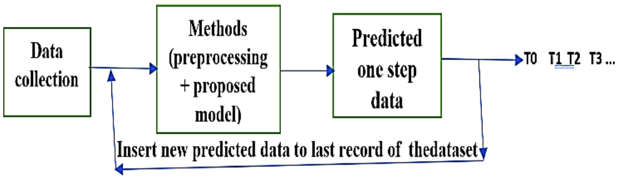


Figure. 3 Proposed method for algorithm 2

Algorithm 1: Proposed Future Human Activity

```
FUNCTION predict_future_activity(model,
input_sequence, steps_ahead):
 predicted_sequence = zeros(shape=(steps_ahead,
input_sequence.shape[1]))
    FOR i = 0 TO steps_ahead-1:
yhat = model.predict(input_sequence.reshape(1,
input_sequence.shape[0],
input_sequence.shape[1]))[0]
predicted_sequence[i] = yhat
input_sequence =
concatenate([input_sequence[1:], yhat.reshape(1,
-1)], axis=0)
RETURN predicted_sequence
predicted_sequence =
predict_future_activity(model, input_sequence,
steps_ahead)
predicted_label =
le.inverse_transform([argmax(predicted_sequence
[0][-1])])

PRINT("Predicted activity:", predicted_label[0])
```

Fig. 2 displays a robust real-time suggested system capable of predicting future human activity using hidden state h(t), which is updated at each time step t based on the input x(t).

Algorithm 1 is known as "predict_future_activity." It takes three primary inputs: a pre-trained LSTM model, an input sequence, and the desired number of steps for future prediction. The primary objective of this function is to generate a sequence comprising predicted future activity levels as its output. The algorithm initializes by creating an empty array labelled "predicted_sequence," which is dimensioned to align with the output specifications of the LSTM layer. It then employs the LSTM layer's output to extract localized features, considering the parameters "steps_ahead" and "num_features." Subsequently, the algorithm utilizes a for loop to predict the forthcoming activity level for each step into the set of internal hidden state variables, including cell hidden

state. The LSTM unit in the layer contains a state c(t) and future, and the previous state h(t-1). One-step prediction assigned to algorithm 2. Each T_0 in Fig. 3 output a one-step prediction added to the last record in the original dataset.

The predicted activity is added to the predicted sequence and the input sequence is shifted by one step and appended with the predicted activity in Fig. 3. After the predict_future_activity function is called and the predicted_sequence is obtained, the predicted label is extracted from the first step of the predicted sequence using argmax and inverse_transform. Finally, the predicted label is printed on the console. Algorithm 2 represents future prediction LSTM represented in Fig. 2 and results in Table 4, below depending on sequences of predicted activity to be added (updated) dataset after the last records, Evaluation can be done with the database by applying the same proposed method.

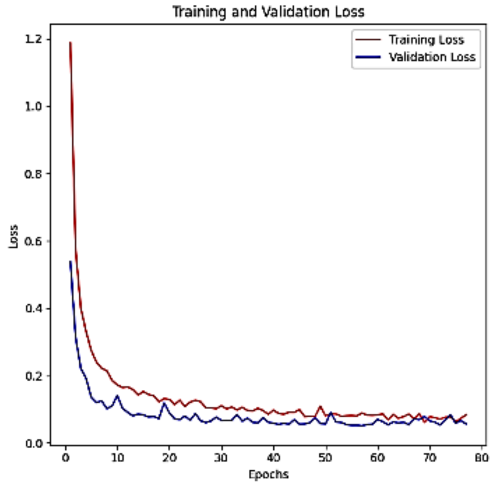
Received: September 2, 2023. Revised: October 10, 2023.

Algorithm 2: Proposed n-steps human activity

function predict_future_activity(model, input_sequence, steps_ahead):
1. For i in range(steps_ahead):
 a. Predict the next activity level by calling model.predict() on the input sequence.
 b. Append the predicted activity to the end of the input sequence using np. append().
c. Remove first element of the input sequence using np. delete() to shift the input sequence by one step.
 2. Return the final predicted sequence.
main program:
1. Load the trained and test set
2. Select the first row of the test set (X_test) and reshape it into a 2D array with shape (1, window_size, num_features).
 3. Set the values of steps_ahead.
4. Call predict_future_activity with the model.
5. Convert the predicted sequence to the original labels
7. Create a new input sequence X_new with a range of values from 0.5 to 1.6.
8. Reshape X_new into a 3D array with shape (1, X_new.shape[0], 1).
9. Predict the activity level for the next day by calling model.predict() on X_new.
10. Print the predicted activity level.

4. Experimental results

Paper's results prove the robustness of the model, as detailed in the methodology and system structure section. The model's primary objective is to emulate human logic, and in this regard, by showing an evaluation involving two distinct methods on the UCI-HAR dataset to validate our approach. To determine the model's efficiency and generalizability, we subjected it to rigorous testing on three additional datasets: Mhealth, WISDM, and PAMP2. These datasets were acquired using consistent data collection procedures to ensure comparability. Our data pre-processing involved the utilization of wavelet transformation to effectively handle time series data. For multiclass classification, a linear encoder was employed. The ensuing Figs. 4, 5, 6, and 7, comprising aspects (a) and (b), present key performance metrics, namely loss, validation loss, accuracy, and validation accuracy, providing a visual representation of our model's performance across various datasets and evaluation criteria, thus enhancing the recognition of different activities and deepening the comprehension of human movement patterns, with the following parameters: LSTM with 32 units, Conv1D= 32 filters, kernel size 3, and ReLU activation, MaxPooling1D=2,Dense= 64 units and ReLU activation, Dropout=0.5, Epochs= 100, batch_size=128, n_samples equal to Number of samples in the input data, n_channels equal to Number of input data channels (features).



(a)

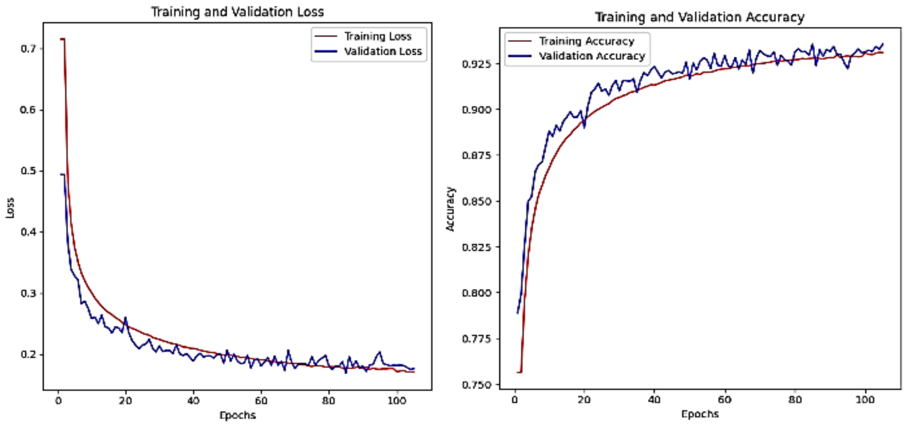
(b)

Figure. 4: (a) UCI-HAR (loss, val_loss) and (b) UCI-HAR (accuracy, val_accuracy)

Methods based UCI-HAR dataset's previous accuracy in [20-24]: 93.18, 95.49, 95.58, and 95.4 respectively. The PAMP consists of data collected from wearable sensors during various physical activities. Table 2 contains the results-based UCI-HAR for implementing Algorithm 2 with three various metrics.

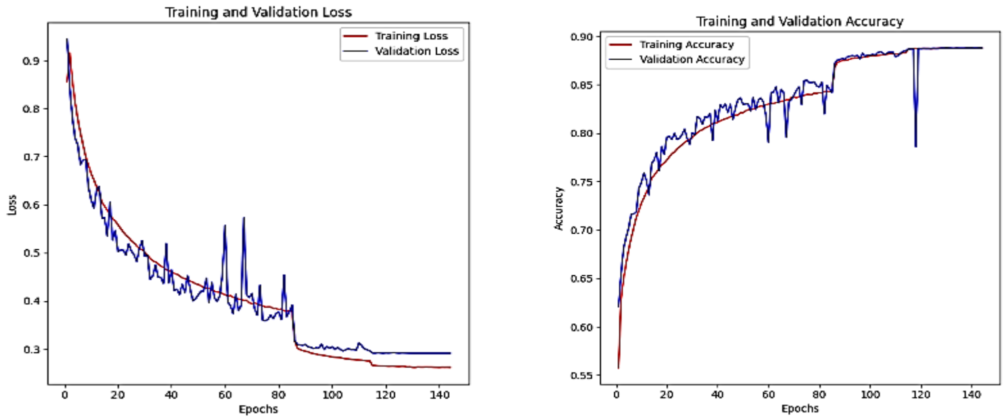
In Table 2 Precision: 0.9861, Recall: 0.9861, and F1-score: 0.9861 were evaluated based on UCI-HAR, epochs 22 which appears high scores compared to results of the state-of-the-art in Table 4 by linking the physical sequence of activities with the temporal-spatial steps involved in them.

Received: September 2, 2023. Revised: October 10, 2023.



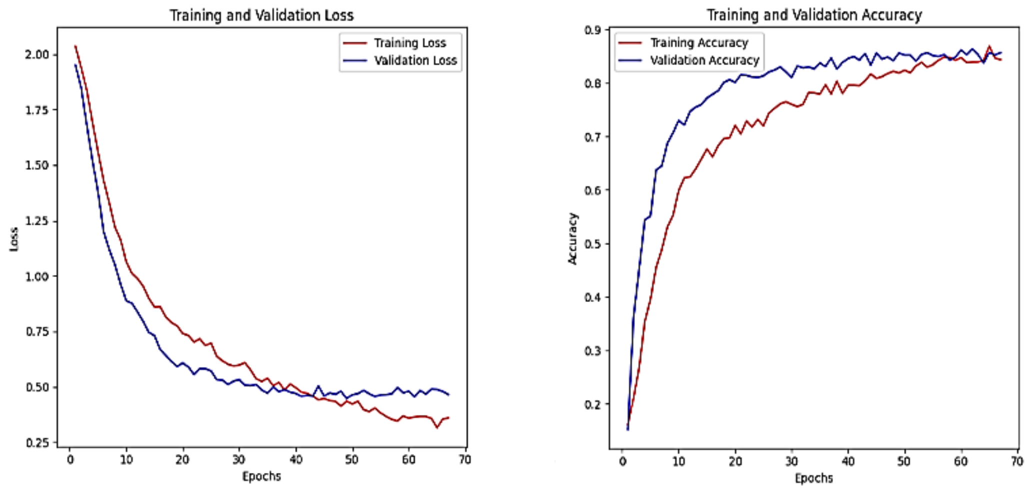
(a) (b)

Figure 5: (a) Mhealth (loss, val_loss) and (b)Mhealth (accuracy, val_accuracy)



(a) (b)

Figure. 6: (a) WISDM (loss, val_loss) and (b) WISDM V2.0 (accuracy, val_accuracy)



(a) (b)

Figure. 7: (a) Pamp2 datasets (accuracy, val_accuracy) and (b) Pamp2 datasets (accuracy, val_accuracy)

Table 3 shows the results of using algorithm 1 to predict future human activity. Table 4 contains results showing the comparison based on UCI-HAR.

Table 5 presents a comparative analysis of the most recent and closely related articles, methods with accuracy referenced as supervised [28], Pseudo label [29], temporal ensemble [30], auto encoder [31], relation prediction [32], and semi-supervised [33] respectively, the datasets, methods, and corresponding test results in the table as follow:

Primarily Table 5 focuses on semi-supervised learning [33] as the last paper published, also the best results compared with the state of the arts, one main

Table 2 Results of proposed algorithm 2

Metrics	Results
F1-score	0.9861
Test Accuracy	0.9892
Confusion Matrix generated	[[703 0 0 0 0 0] [0 626 2 0 0 0] [0 2 576 0 0 0] [0 0 0 701 10 0] [0 0 0 45 735 0] [0 0 0 0 0 779]]

Table 3 Results of proposed methods

Dataset	Method	Accuracy
UCI-HAR Epoch 77	Modified (DWT) with LSTM+ CON1D	Accuracy: 0.9852, figure 4 shows evaluation metrics
Mhealth Epoch 88		Accuracy: 0.9321, figure 5 shows evaluation metrics
WISDM Epoch 162		Accuracy: 0.8943, Figure6 shows evaluation metrics
PAMAP2 Epoch 79		Accuracy: 86.85% and Figure 7 shows evaluation metrics

Table 4 Comparison models based UCI-HAR

Model	Year	Accu-racy	Difference point(s)
BiLSTM [25]	2023	**95.79**	1- Higher complexity o(n^2) 2- Dropout = 0.3, 0.2 and L2 regularizer = 1e-2.
Bi-LSTM [26]	2023	**97.96**	1- Higher complexity o(n^2) 2- Only five daily activities 3- Predict future activity.
Temporal Conv-LSTM [27]	2022	**91.6**	1-A hybrid architecture was employed, with parallel feature learning pipelines. 2-High complexity o(n^2) 3- sliding and size window.
Our model		**98.92**	1-Less complexity (single layer O(n) LSTM and CONV1D) 2- Optimizer: Adam, Batch size:64, lr:0.1, training data 0.7 and 0.8.

strength of our same methodology applied to all datasets, is less complexity due to the use of multiple methods to build framework. In the case of the mHealth dataset, study [33] suggests re-categorizing label data into six classes to address the challenges posed by complicated activities while in our work all datasets were pre-processed without changing.

5. Conclusion:

This paper utilizes comprehensive visual characteristics and principal considerations to

Table 5 Comparison of varying methods

Dataset	Method	20 % Test set
UCI-HAR Epochs=200	Supervised	91.8
	Pseudo label	92.7
	Temporal Ensemble	93.2
	Auto Encoder	91.5
	Relation Prediction	92.4
	Semi-supervised	93.5
	Our method	**0.9903**
WISDM	Supervised	83.4
	Pseudo label	86.1
	Temporal Ensemble	87.1
	Auto Encoder	84.9
	Relation Prediction	86.0
	Semi-supervised	88.1
	Our method	**92.46**
PAMAP2 Epochs=55	Supervised	83.2
	Pseudo label	84.2
	Temporal Ensemble	83.8
	Auto Encoder	84.2
	Relation Prediction	84.6
	Semi-supervised	85.4
	Our method	**0.8552**
mHealth epochs=89	Supervised	95.6
	Pseudo label	95.8
	Temporal Ensemble	94.5
	Auto Encoder	96.3
	Relation Prediction	95.8
	Semi-supervised for 6 classes out of 12	97.5
	Our method	**94.48**

concurrently predict individuals' paths and activities by presenting two different methodologies with a new procedure for applying a windowing function to the (WT), first algorithm is a novel proposed framework that implies temporal-spatial with the LSTM-CONV1D model is (DL) for handling sequences of activities and validated via four different datasets to satisfy supervision from heterogenous based sensor sources. Second, are presented in this work, the predict_future_activity function can be used to make predictions for future time steps based on a given input sequence based on the UCI-HAR.

Future work planning to apply the proposed method to other relevant datasets, following the same methodology phases outlined. The main finding of the proposed framework can be efficiently used for the smarthome and fall system.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

Conceptualization, Mohammad and Ali; methodology, Mohammad; software, Ali; validation, Mohammad; formal analysis, Mohammad and Ali; investigation, Mohammad; resources, Mohammad; data curation, Mohammad writing-original draft preparation, Mohammad; writing-review and editing, Mohammad and Ali; visualization, Ali; supervision, Mohammad; project administration, Mohammad; funding acquisition, Mohammad. All authors have read and approved the final manuscript.

References

[1] V. B. Semwal, N. Gaud, P. Lalwani, V. Bijalwan, and A. K. Alok, "Pattern identification of different human joints for different human walking styles using inertial measurement unit (IMU) sensor", *Artificial Intelligence Review*, Vol. 55, pp. 1149-1169, 2022.

[2] Jain, V. B. Semwal, and P. Kaushik, "Stride segmentation of inertial sensor data using statistical methods for different walking activities", *Robotica*, Vol. 40, pp. 2567-2580, 2022.

[3] R. Jain, V. B. Semwal, and P. Kaushik, "Deep ensemble learning approach for lower extremity activity recognition using wearable sensors", *Expert Systems*, Vol. 39, p. e12743, 2022.

[4] S. Mekruksavanich and J. Jitpattanakul, "Biometric User Identification Based on Human Activity Recognition Using Wearable Sensors: An Experiment Using Deep Learning Models", *Electronics*, Vol. 10, No. 3, p. 308, 2021.

[5] T. H. Tan, J. Y. Wu, S. H. Liu, and M. Gochoo, "Human Activity Recognition Using an Ensemble Learning Algorithm with Smartphone Sensor Data", *Electronics*, Vol. 11, No. 3, p. 322, 2022.

[6] E. Hysenllari, J. Ottenbacher, and D. McLennan, "Validation of human activity recognition using a convolutional neural network on accelerometer and gyroscope data", *German Journal of Exercise and Sport Research*, Vol. 52, pp. 248-252, 2022.

[7] N. Dua, S. N. Singh, and V. B. Semwal, "Multi-input CNN-GRU based human activity recognition using wearable sensors", *Computing*, Vol. 103, pp. 1461-1478, 2021.

[8] Q. Zou, Y. Wang, Q. Wang, Y. Zhao, and Q. Li, "Deep learning-based gait recognition using smartphones in the wild", *IEEE Transactions on Information Forensics and Security*, Vol. 15, pp. 3197-3212, 2020.

[9] A. Prasad, A. K. Tyagi, M. M. Althobaiti, A. Almulihi, R. F. Mansour, and A. M. Mahmoud, "Human Activity Recognition Using Cell Phone-Based Accelerometer and Convolutional Neural Network", *Applied Sciences*, Vol. 11, No. 3, p. 12099, 2021.

[10] Vasanthanageswari and P. Prabhu, "IMPROVING SVM CLASSIFIER MODEL USING TREE STRUCTURED PARZEN ESTIMATOR OPTIMIZATION FOR CROP PREDICTION", *Journal of Theoretical and Applied Information Technology*, Vol. 100, pp. 6808-6818, 2022.

[11] Y. Tang, L. Zhang, Q. Teng, F. Min, and A. Song, "Triple cross-domain attention on human activity recognition using wearable sensors", *IEEE Transactions on Emerging Topics in Computational Intelligence*, Vol. 6, pp. 1167-1176, 2022.

[12] M. Sivakami and P. Prabhu, "Classification of Algorithms Supported Factual Knowledge Recovery from Cardiac Data Set", *International Journal of Current Research and Review*, Vol. 13, Issue 6, pp. 161-166, March 2021.

[13] R. Chetan, D. V. Ashoka, and B. V. Ajay Prakash, "IMLAPC: interfused machine learning approach for prediction of crops", *Reviews in Intelligent Automation*, Vol. 36, pp. 169-174, 2022.

[14] C. Han, L. Zhang, Y. Tang, W. Huang, F. Min, and J. He, "Human activity recognition using wearable sensors by heterogeneous convolutional neural networks", *Expert Systems with Applications*, Vol. 198, p. 116764, 2022.

[15] S. Miao, L. Chen, R. Hu, and Y. Luo, "Towards a dynamic inter-sensor correlations learning framework for multi-sensor-based wearable human activity recognition", In: *Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, pp. 1–25, 2022.

[16] W. N. Ismail, H. A. Alsalamah, M. M. Hassan, and E. Mohamed, "AUTO-HAR: An adaptive human activity recognition framework using an automated CNN architecture design", *Heliyon*, Vol. 9, Issue 2, p. e13636, 2023.

[17] M. A. Khatun, M. A. Yousuf, S. Ahmed, M. Z. Uddin, S. A. Alyami, S. A. Ashhab, H. F. Akhdar, A. Khan, A. Azad, and M. A. Moni, "Deep CNN-LSTM With Self-Attention Model for Human Activity Recognition Using Wearable Sensor", *IEEE Journal of Translational Engineering in Health and Medicine*, Vol. 10, pp. 1-16, Art No. 2700316, 2022.

[18] A. Tehrani, M. Y. Tabari, A. Z. Salmasi, and R. Enayatifar, “Wearable Sensor-Based Human Activity Recognition System Employing Bi-LSTM Algorithm”, *The Computer Journal, Bxad035*, 2023.

[19] O. Nafea, W. Abdul, G. Muhammad, and M. Alsulaiman, “Sensor-Based Human Activity Recognition with Spatio-Temporal DL”, *Sensors*, Vol. 21, No. 6, p. 2141, Mar. 2021.

[20] Y. H. Lin and J. Wu, “A Novel Multichannel Dilated Convolution Neural Network for Human Activity Recognition”, *Mathematical Problems in Engineering*, Vol. 2020, p. e5426532, Jul. 2020.

[21] S. Wang and X. Zhu, “A hybrid deep neural networks for sensor-based human activity recognition”, In: *Proc. of IEEE/2020 12th International Conference on Advanced Computational Intelligence (ICACI),* Dali, China, pp. 486–491, 2020.

[22] M. G. Ragab, S. J. Abdulkadir, and N. Aziz, “Random search one-dimensional CNN for human activity recognition”, In: *Proc. of IEEE International Conference on Computational Intelligence, ICCI*, pp. 86–91, 2020.

[23] Z. Salem and A. P. Weiss, “Improved Spatiotemporal Framework for Human Activity Recognition in Smart Environment”, *Sensors*, Vol. 23, p. 132, 2023.

[24] A. Sadeghian, A. Alahi, and S. Savarese, “Tracking the untrackable: Learning to track multiple cues with long-term dependencies”, In: *Proc. of the IEEE International Conference on Computer Vision*, pp. 300–311, 2017.

[25] A. Tehrani, M. Y. Tabari, A. Z. Salmasi, and R. Enayatifar, “Wearable Sensor-Based Human Activity Recognition System Employing Bi-LSTM Algorithm”, *The Computer Journal, Bxad035*, 2023.

[26] I. E. Jaramillo, C. Chola, J. Y. Jeong, J. H. Oh, H. Jung, J. H. Lee, W. H. Lee, and T. S. Kim, “Human Activity Prediction Based on Forecasted IMU Activity Signals by Sequence-to-Sequence Deep Neural Networks”, *Sensors*, Vol. 23, p. 6491, 2023.

[27] M. H. M. Noor, S. Y. Tan, and M. N. A. Wahab, “Deep Temporal Conv-LSTM for Activity Recognition”, *Neural Process Letters*, Vol. 54, pp. 4027–4049, 2022.

[28] M. A. Basset, H. Hawash, V. Chang, R. Ranjan, and M. Niranjan, “Deep learning for heterogeneous human activity recognition in complex IoT applications”, *IEEE Internet of Things Journal*, Vol. 9, No. 8, pp. 5653–5665, 2022.

[29] D. H. Lee, “Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks”, In: *Proc. of Workshop on challenges in representation learning, International Conference on Machine Learning*, p. 896, 2013.

[30] S. Laine and T. Aila, “Temporal ensembling for semi-supervised learning”, In: *Proc. of 5th International Conference on Learning Representations*, 2017.

[31] M. Zeng, T. Yu, X. Wang, L. T. Nguyen, O. E. Gundersen, and I. R. Lane, “Semi-supervised convolutional neural networks for human activity recognition”, In: *Proc. of 2017 IEEE International Conference on Big Data*, pp. 522–529, 2017.

[32] H. Fan, F. Zhang, R. Wang, X. Huang, and Z. Li, “Semi-supervised time series classification by temporal relation prediction”, In: *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3545–3549, 2021.

[33] Y. Qu, Y. Tang, X. Yang, Y. Wen, and W. Zhang, “Context-aware mutual learning for semi-supervised human activity recognition using wearable sensors”, *Expert Systems with Applications*, Vol. 219, p. 119679, 2023.