



An optimized hybrid deep learning model using ensemble learning approach for human walking activities recognition

Vijay Bhaskar Semwal¹ · Anjali Gupta¹ · Praveen Lalwani²

Accepted: 22 March 2021 / Published online: 6 April 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Recent advancements in edge computing devices motivate us to develop a sustainable and reliable technique for multiple gait activities recognition using wearable sensors. This research work presents the multitask human walking activities recognition using human gait patterns. Human locomotion is defined as the change in the joint angles of hip, knee and ankle. To achieve the aforementioned objective, the data are collected for 50 subjects in a controlled laboratory environment using inertial measurement unit (IMU) sensors for 7 different activities. The IMU sensor is placed on the chest, left thigh, and right thigh. Total 100 samples are collected for all 7 activities. The sampling rate considered was 50 Hz. Following 7 walking activities are performed for all the 50 subjects: (i) natural walk, (ii) standing, (iii) climbing stairs, (iv) cycling, (v) jogging, (vi) running, (vii) knees bending (Crouching). The major contribution of this research paper is the design of four hybrid deep learning models to provide the generic activity recognition framework and tune the performance. The following combination of the deep learning model is designed for the classification of gait activities, namely, convolution neural network–long short-term memory (CNN–LSTM), CNN–gated recurrent unit (CNN–GRU), LSTM–CNN and LSTM–GRU. To support edge computing, the ensemble learning is utilized to optimized the model size. The proposed ensemble learning-based hybrid deep learning framework has provided a promising classification accuracy of 99.34% over other models. The other models namely CNN, LSTM, GRU, CNN–LSTM, LSTM–CNN, CNN–GRU, GRU–CNN have achieved 97.26%, 90.67%, 77.38%, 97.83%, 94.35%, 97.64%, 96.98% accuracy, respectively, on our HAG data set. The proposed technique is also validated on MHEALTH data set for comparative analysis. The hybrid deep learning model in combination with ensemble learning has outperformed other techniques. The optimized code can be used on small computation devices for walking activity recognition.

The work is Funded by SERB, DST, Govt. of India to Dr. Vijay Bhaskar Semwal under Early Career Award (ECR) with DST No: ECR/2018/000203 dated 04-June-2019.

Extended author information available on the last page of the article

Keywords Ensemble learning · Activity recognition · Wearable sensor · Gait analysis · Hybrid deep learning · Human robot interaction(HRI)

1 Introduction

1.1 Overview

A bipedal walking robot is a kind of humanoid robot, which is very similar to human morphology. It mimics human behavior and is devised to perform human-specific tasks. Currently, humanoid robots are not capable to walk properly like human beings. The human location is referred by the term human gait. It is a manifestation of the change in the joint angles of hip, knee and ankle. The human gait is considered a very unique biometric pattern of human being [1, 2]. Compare to other biometrics, it is unobstructive, difficult to spoof, and permanent. However, it is a very complex learning process which involves interaction among different parts of the brain along with spinal cord, motor action and muscle interaction [3]. It is highly variable and depends on several parameters, e.g., speed, cloth, walking style, mental health condition, etc. [4]. It is affected by several factors like mental status, fatigue situation and cognitive status. These are measured using changes in the joints ankle value, step length, stride length, ground clearance, etc. It is potentially being used in measuring the various health issues like cognitive decline due to aging, freezing of gait, Parkinson, etc. It is one of the complicated physical activity which involves the coordination and synchronization of the various body parts together [5]. The gait is widely used for clinical assessment of neurological disorder patients, early diagnosis of gait abnormality, etc [6, 7]. It is also used for robotics walk generation and enhancing the cognitive capability to enable HRI [8, 9] and haptic [10]. Due to its biphasic and bipedal nature, it is popular for various human activities recognition [11].

1.2 Different sub-phases and events of gait cycles

The gait cycle consists of seven sub-phases as shown in Fig. 1. Each sub-phase corresponds to some percentage of gait cycle as shown in Table 1. The stance phase also known as double support phase(DSP), when both the feet touch the ground. On the other hand, swing phase also known as single support phase(SSP) as one limb touches the ground, while the other is off the ground to begin the next step [12]. Figure 2 represents the different event and configuration of walking sub-phases.

1.3 Author's contribution

Highlights of author's contribution as follows:

1. *Database creation and analysis* Total 50 subjects data are collected for seven walking activities. The data are collected using tri-axial 9-Degree of freedom(Dof)

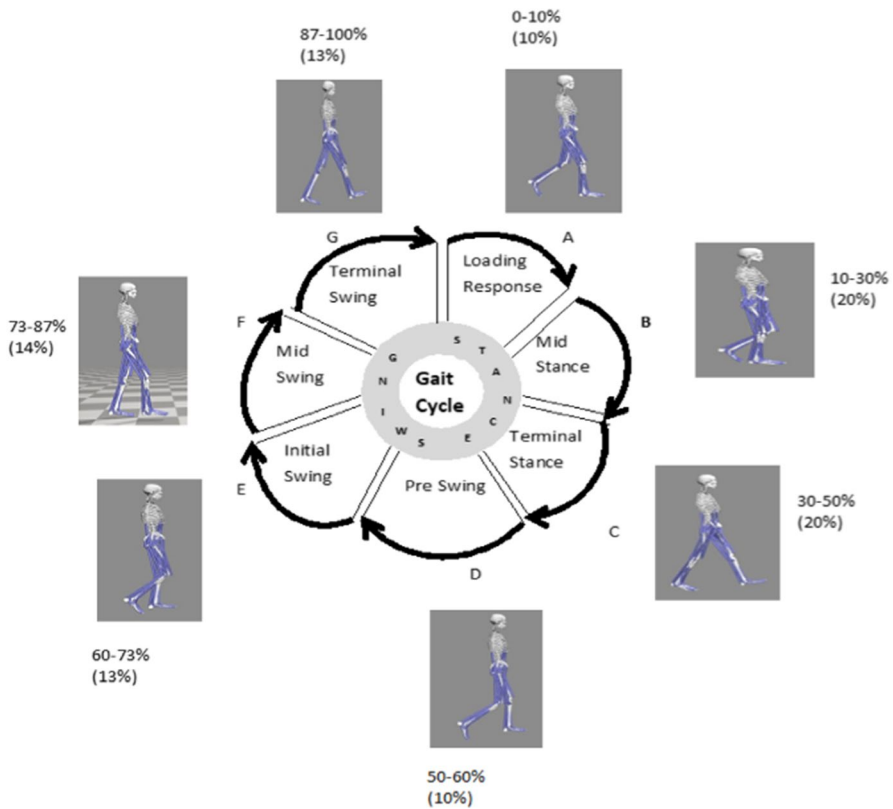
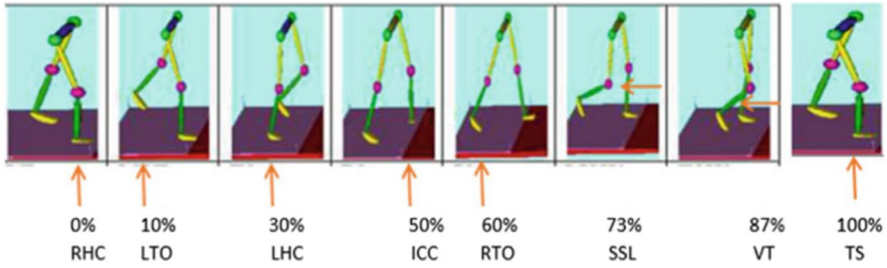


Fig. 1 Different walking sub-phases

IMU sensor BWT61CL MPU6050. The IMU sensors were placed on chest, left thigh and right thigh for three different modality of data set.

2. *Data representation* The sampling rate was 50 Hz, and each activity is recorded for 100 samples. Total 24 attributes are collected from sensors for each sample. The sensor used was triaxial acceleration, gyroscope, and magnetometer placed at chest, left, and right thigh. The tensor size was $100 \times (50 \times 24)$.
3. *Data pre-processing* The extended Kalman filter is used for removable of noise during sensor reading.
4. *Modelling of deep learning and hybrid deep learning classifiers* To classify the data of different activities, four different hybrid deep learning models, namely, CNN-LSTM, CNN-GRU, LSTM-CNN, and LSTM-GRU are applied. To provide the generic solution, the ensemble learning method was also applied.
5. *Performance analysis* The proposed method has achieved 99.34% accuracy which is quite high and accurate. The MHEALTH data set is used for verification of model as benchmark data set.



STANCE(60%)				SWING (40%)		
Loading Response [0-10] %	Mid Stance [10-30%]	Terminal Stance [30-50%]	Pre Swing [50-60%]	Initial Swing [60-73%]	Mid Swing [73-87%]	Terminal Swing [87-100%]
Double Support	Right Single Support		Double Support	Left Single Support		

RHC - Right heel contact , LTO - Left toe off, LHC- Left heel contact, ICC- Initial contact of the contralateral limb, RTO – Right toe off, SSL - Swimming limbs opposite to stance limb,VT- Vertical tibia, TS- Terminal swing.

Fig. 2 Different gait sub-phases event

Table 1 Percentage wise division of different walking sub-phases

S. No.	Walking sub-phase	Percentage range
<i>Stance phase</i>		
1	Initial contact (IC)	[0–2]%
2	Loading response (LR)	[2–10]%
3	Mid stance (MST)	[10–30]%
4	Terminal stance (TS)	[30–50]%
<i>Swing phase</i>		
5	Pre swing(PS)	[50–60]%
6	Initial swing(IS)	[60–73]%
7	Mid swing(MSW)	[73–87]%
8	Terminal swing(TSW)	[87–100]%

1.4 Novelty

The paper has extensively explored the 7 walking activities data recorded through IMU sensor with extended Kalman filter in a controlled laboratory environment for 50 subjects. The paper has also opened the direction of how human activities can be used for HRI, building cognitive capability, diagnosis of disease at early stage, and study of human adaptive locomotion. The novelty of this research work is multidimensional. It starts from data collection to selection of filter and finally designing of hybrid deep learning-based model. The generic activity recognition framework is designed using 4 hybrid deep learning models namely CNN–LSTM, LSTM–CNN,

CNN–GRU, and GRU–LSTM for all 7 walking activities. Finally, to provide the visual presentation of results the performance matrix, classification matrix and accuracy curves are included. To provide the generic solution, the ensemble learning-based method is employed.

1.5 Organization of the paper

The rest of the paper is structured as follows. The second section includes a literature review of related work about human activity recognition using a wearable sensors. It also describes several deep learning models with ensemble learning for walking activity data recognition. The third section is preliminaries. It provides the detail of all prerequisites require for study including notation and abbreviation, data pre-processing filters and different deep learning techniques. Section 4 contains a methodology and the proposed algorithm for the activities recognition system. This section illustrates data collection procedures in a controlled laboratory environment and pre-processing of raw data. This section also describes the data set and different human walking activity recognition algorithms. Section 5 discusses results and analysis of different activity classification using hybrid deep neural network models namely, CNN, CNN–LSTM, LSTM–CNN, CNN–GRU, and GRU–CNN. The last Sect. 6 incorporates conclusion and proposed future research work.

2 Literature review of related works

The gait is very commonly used in human activity recognition. Hsu et al. [13] used multiple wearable sensors for analyzing and classifying the gait of patients with neurological disorders like multiple sclerosis, cerebral palsy and stroke patients. Another novel approach was proposed by Mekruksavanich et al. [14] where smartphones as wearable sensors were used to collect data. The study was made to classify gait patterns for three different activities like walking upstairs, walking downstairs and walking on floor using LSTM. Jennifer et al. [15] have collected 7 different human activities data using wearable sensor and classified using DNN. Ioannis et al. [16] presented work for classification of neurological gait disorders using multitask feature learning. Semwal et al. [17] have proposed human gait state prediction using cellular automata and classification using ELM. Semwal et al. [18] have provided the optimized feature based on incremental feature analysis for gait data classification. In subsequent work, Semwal et al. [19] have also designed the vector field for gait sub-phases for reconstruction of gait. Chen et al. [20] have proposed deep convolutions neural networks based on multistatic micro-Doppler signatures for gait classification. Semwal et al. [21] proposed the restricted Boltzmann machine-based DNN model for gait activities classification. Poschadel et al. 2017 used dictionary learning method for gait classification [22, 23].

Some deep learning approaches like CNN allow classification directly on raw data without performing features extraction [24]. The CNN model performs three important work, i.e., sparse interaction, parameter sharing and equivalent representation.

After convolution pooling and fully connected layers are used for classification and regression tasks [25].

Gupta et al. [26] have proposed a method for human activity recognition using ensemble learning technique. The results proved that the ensemble learning has outperformed among various techniques used for activity recognition and classification. For dimensionality reduction, principal component analysis method was used. The limitation was that the proposed algorithm was implemented on standard data sets and autoencoders could also be used for dimensionality reduction. Wang et al. [27] performed gait classification using ensemble CNN model which also suffered from several disadvantages like it could not use heterogeneous CNN classifiers for classification. The concept of integration of different classifiers was also proposed by Sun et al. [28] in 2020 on multigait feature for human gait classification. The ensemble learning method enhanced the overall recognition accuracy of the proposed framework but it also increased the complexity of the network. Wang et al. [29] have proposed a method for gait recognition which was invariant to viewing directions and incorporated heterogeneous techniques combined together for learning. However, the results were obtained from experimental data sets which could not further be extended in practical application. Nandi et al. [9] have proposed the hybrid automata-based model for generation of human locomotion trajectories in 2016. These model proved to be very efficient in calculating the joint trajectories correctly and this work could also be extended to push recovery analysis further.

3 Preliminaries

This section presents the description of notations and abbreviations, Kalman filter for data pre-processing and noise removal, benchmark data set description and various deep learning algorithms used in proposed methodology.

3.1 Notations and abbreviations

This subsection presents the notation and abbreviation used in the proposed methodology. Table 2 describes all the notation.

3.2 Standard filter for IMU data

3.2.1 Extended Kalman filter

The noisy measurement of data is removed using the extended Kalman filter. The noisy reading is due to environment noise, self-occlusions, loss of accuracy due to fast movements, etc., during the data acquisition. Likelihood function is used for estimation of the variance of the noise processing the capturing of the raw data using IMU sensor. Extended Kalman filter is applied to remove the noise from the model [30]. Extended Kalman filter has the ability to smooth the accelerometer

Table 2 Notation and abbreviations

Notation	Description
ε	Noise factor
$\mathcal{N}(0, Q)$	Gaussian distribution with 0 mean
δ	Noise factor
z_t	Final layer output
h_t	Hidden layer output
TP	True positive
TN	True negative
FP	False positive
FN	False negative
β	Weight coefficient

and gyroscope reading. The extended Kalman filter performs the nonlinear transformations also.

$$x_{t+1} = Ax_t + \varepsilon, \quad \text{where } \varepsilon \sim \mathcal{N}(0, Q) \quad (1)$$

$$y_{t+1} = Ax_{t+1} + \delta, \quad \text{where } \delta \sim \mathcal{N}(0, R) \quad (2)$$

where, x represents state vector, y represents measurement vector and t represents the discrete time index. To model nonlinear transformations, extended Kalman filter is used:

$$x_{t+1} = f(x_t) + \varepsilon, \quad \text{where } \varepsilon \sim \mathcal{N}(0, Q) \quad (3)$$

$$y_{t+1} = g(x_{t+1}) + \delta, \quad \text{where } \delta \sim \mathcal{N}(0, R) \quad (4)$$

Nonlinearity in the transformations is dealt with the usage of Taylor expansion. In the presented case, the transformations are approximated using the first-order Taylor expansion evaluated at the mean estimate of x at time t , μ_t :

$$x_{t+1} \approx f(\mu_t) + \frac{\partial f}{\partial x}(\mu_t)(x_t - \mu_t) + \varepsilon \quad (5)$$

$$y_{t+1} \approx g(\mu_{t+1}) + \frac{\partial g}{\partial x}(\mu_{t+1})(x_{t+1} - \mu_{t+1}) + \delta \quad (6)$$

The first-order derivative term $\frac{\partial f}{\partial x}(\mu_t)$, also referred to as the Jacobian, is used here as the linear transformation, while the $f(\mu_t)$ term just serves to shift the mean of the transformation.

Extended Kalman filter also take care to include limiting conditions based on the joints displacements per frame.

The update equations obtained for our mean and variance estimates for the motion model using the nonlinear transformation above, we get the following new update equations for our mean and variance estimates for the motion model:

$$\mu_{t+1}^- = f(\mu_t^+) \tag{7}$$

$$\Sigma_{t+1}^- = J_f \Sigma_t^+ J_f^T + Q \tag{8}$$

and for the sensor model:

$$\mu_{t+1}^+ = \mu_{t+1}^- + \Sigma_{t+1}^- J_g^T (J_g \Sigma_{t+1}^- J_g^T + R)^{-1} [y_{t+1} - g(\mu_{t+1}^-)] \tag{9}$$

$$\Sigma_{t+1}^+ = \Sigma_{t+1}^- - \Sigma_{t+1}^- J_g^T (J_g \Sigma_{t+1}^- J_g^T + R)^{-1} J_g \Sigma_{t+1}^- \tag{10}$$

where $J_f = \frac{\partial f}{\partial x}(\mu_t^+)$ and $J_g = \frac{\partial g}{\partial x}(\mu_{t+1}^-)$.

3.3 Description of deep learning algorithms

3.4 Convolution neural network (CNN)

The designing of CNN/ConvNet is inspired from the visual cortex of the animal. ConvNet are unsupervised algorithms so they have the capability to learn from the filters and examples. ConvNet is a mathematical structure whose building blocks are convolution, max pooling/average pooling and dense layers. The first two components that is convolution and pooling are responsible for feature extraction and the dense layer is used to map these features to the final output. In order to improve the performance of the model, the depth of the network also increases which comes up with the problem of additional cost. So the CNN model becomes expensive with deeper layers. Another problem associated with the deep layer architecture is overfitting. The problem of overfitting and high computational cost due to deeper networks was overcome by hybrid ensemble learning model. Equation 11 represents the convolution operation.

$$(f * g)(z_1, z_2) = \sum_{\substack{(x_1 + y_1 = z_1) \\ (x_2 + y_2 = z_2)}} f(x_1, x_2) \cdot g(y_1, y_2) \tag{11}$$

where f is data and g is the kernel.

3.4.1 Dropout

In dense layer, problem of co-adaption occurs. Co-adaption means when more than one neurons in a layer extract the or similar information from the input. This problem

majorly occurs when connection weights attached to few neurons are same. This results in the problem of drainage of resources in obtaining same information and leading to the problem of overfitting. Dropout technique is used to solve the problem of co-adaptation in neural network. In this technique, randomly selected neurons are dropped by turning their neuron values to zero in each layer during the training process. The rate at which these neurons are turned off or dropped is known as dropout rate. Dropout also improves the learning rate of a model. Dropout is generally implemented after the fully connected layers in the neural network data sets used.

3.4.2 Convolution neural network–long short-term memory (CNN–LSTM)

It is a hybrid model that tries to combine the powers of convolutional networks followed by long short-term memory (LSTM). CNN is basically involved in feature extraction, and LSTM is responsible for the learning purpose. CNN–LSTM can apply a variety of operations involving sequential or time series inputs and this model is spatially and temporally deep which makes it fit for such kind of classification tasks. Originally it was known as long term recurrent convolutional network (LRCN). Lets have a deeper look into the LSTM. Its working is governed by following equations:

$$\begin{aligned}
 \text{function}_t &= \sigma(\text{weight}_f \cdot [\text{hidden}_{t-1}, \text{input}_t] + \text{bias}_f) \\
 \text{intermediate}_t &= \sigma(\text{weight}_i \cdot [\text{hidden}_{t-1}, \text{input}_t] + \text{bias}_i) \\
 C_{\text{temp}} &= \tanh(\text{weight}_c \cdot [\text{hidden}_{t-1}, \text{input}_t] + \text{bias}_c) \\
 C_t &= \text{function}_t * C_{t-1} + \text{intermediate}_t * c_{\text{temp}} \\
 \text{output}_t &= \sigma(\text{weight}_{\text{output}} \cdot [\text{hidden}_{t-1}, \text{input}_t] + \text{bias}_{\text{output}}) \\
 h_t &= \text{output}_t * \tanh(C_t)
 \end{aligned} \tag{12}$$

Here, we have divided the shape of our data into frames, so it becomes 3-dimensional.

3.4.3 Convolution neural network gated recurrent unit (CNN–GRU)

Gated recurrent units follow a very similar approach to long short-term memory units. GRU has an updated gate and a reset gate which are responsible for the flow of information vectors. These gates combinedly decide what part of the tensor needs to be remembered in the next step and which may be updated. Lets have a closer look on the architecture of Gated recurrent units. The gated recurrent unit is performing the following operation:

$$\begin{aligned}
 z_t &= \sigma(\text{weight}_z \cdot [\text{hidden}_{t-1}, \text{input}_t]) \\
 r_t &= \sigma(\text{weight}_r \cdot [\text{hidden}_{t-1}, \text{input}_t]) \\
 h_{\text{temp}} &= \tanh(\text{weight} \cdot [r_t * h_{t-1}, \text{input}_t]) \\
 h_t &= (1 - z_t) * \text{hidden}_{t-1} + z_t * h_{\text{temp}}
 \end{aligned} \tag{13}$$

Here we have divided the shape of our data into frames, so it becomes 3-dimensional.

3.5 Ensemble learning

Ensemble learning technique used to avoid overfitting by reducing the complexity of model. It used to consider the average performance of different classifier and different hyperparameter. So, if any classifier performs poor and another performs better, the ensemble learning used to combine all results. The average of all classifier will provide the much needed generality of model by reducing variance and model complexity.

4 Proposed methodology

This section consists of data set description and pre-processing, flowchart of proposed architecture, and algorithm.

4.1 Data acquisition system

4.1.1 Data set description

The proposed work has used a multiple-task gait analysis approach using tri axial IMU sensors data. The IMU sensor is placed at chest, left thigh and right thigh. The data were collected for 50 subjects for 7 different activities in controlled laboratory environment. The data set contains data from the various tasks: (i) natural walk, (ii) standing, (iii) climbing stairs, (iv) cycling, (v) jogging, (vi) running, (vii) knees bending (Crouching). Figure 3 shows the data collection of different walking subjects using IMU sensor. The data are collected in controlled laboratory environment. Total 50 subjects data are collected for seven walking activities.

4.1.2 Data set validation

To validate our collected data set, the previously available MHEALTH data set is considered as benchmark data set for comparing results with our collected data [31]. The MHEALTH data set is a physical activity monitoring data set. The mobile Health data set captures different physical activities based on multimodal body sensing. Ten subjects were made to perform 12 different physical activities like standing, walking, running, cycling, jogging, climbing stairs, crouching and so on. The data were captured by making use of 3 inertial measurement units (IMU) sensors and a heart rate monitor. Three IMU sensors were placed on each subject chest, right wrist and left ankle. The number of attributes obtained from this data set are 23. ECG measurements were also used for heart monitoring. This data set is used for Human activity recognition for identifying and classifying different physical activities [32].

4.2 Pre-processing steps: an extended Kalman filter

Due to inherent and physical limitation and high sensitivity, the sensor may record the noisy data. Hence, pre-processing of data is important before providing the deep

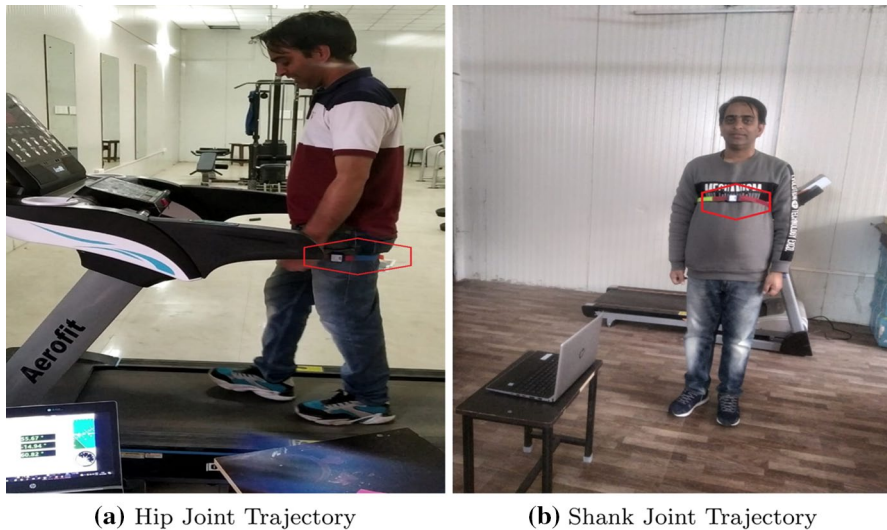


Fig. 3 Different subject(s) data collection using IMU **a** Sensor placed on left and right thigh joint; **b** sensor placed on chest

learning-based classifier. The extended Kalman filter is utilized for smoothing the data. A window frame of 0.8 s was taken with overlapping of 50%.

4.3 Proposed architecture

Figure 4 shows the working flowchart of the proposed activities recognition system.

4.4 Model design and validation

The randomly Shuffling raw data are random splitting in to train, test and validation data. When designing the classification model, it is the best choice to test their accuracy on different data. For this purpose, whole data are divided into three part 60% training, 20% validation and 20% test data set. Further, the validation data are also shuffled with training data using cross-validation.

4.4.1 Data classification and parameter tuning

The data are classified using proposed hybrid deep learning-based classifier. Total 7 classifier is designed for classification purpose namely CNN, LSTM, GRU, CNN–LSTM, LSTM–CNN, CNN–GRU, GRU–CNN AND Ensemble Learning model. Figure 5 shows the proposed hybrid deep learning architecture. Tables 3, 4, 5 and 6 shows the list of hyperparameter used by different classifier.

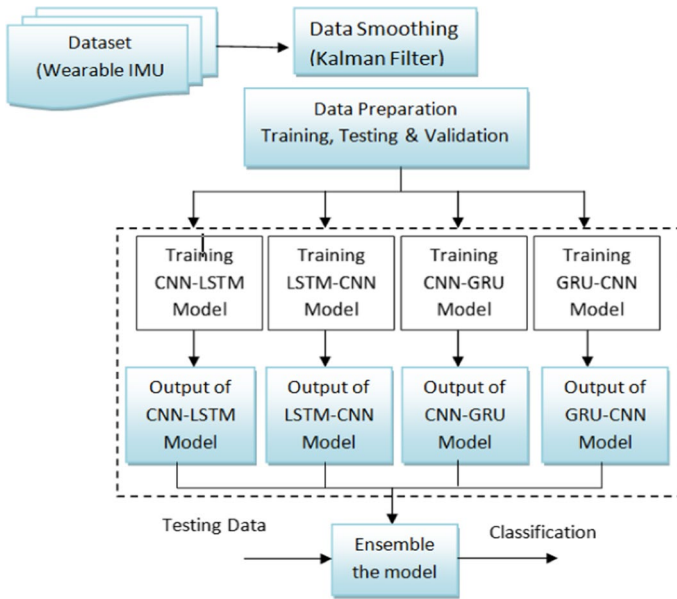


Fig. 4 Working flowchart of the proposed activities recognition system

4.4.2 Fivefolds cross-validation

The fivefold cross-validation is implemented. Figure 6 refers to fivefold cross-validation architecture.

4.5 Performance evaluation parameter

The performance of any classifier used to measure in term of accuracy, precision, f score and recall value. The following section will deal with these parameters.

4.5.1 Accuracy

Accuracy is a performance measure which can be defined as the ratio of correctly predicted observation to the total observation

$$\text{Accuracy} = \frac{c}{n} \tag{14}$$

Where c is correctly classified activities and n is the total number of activities.

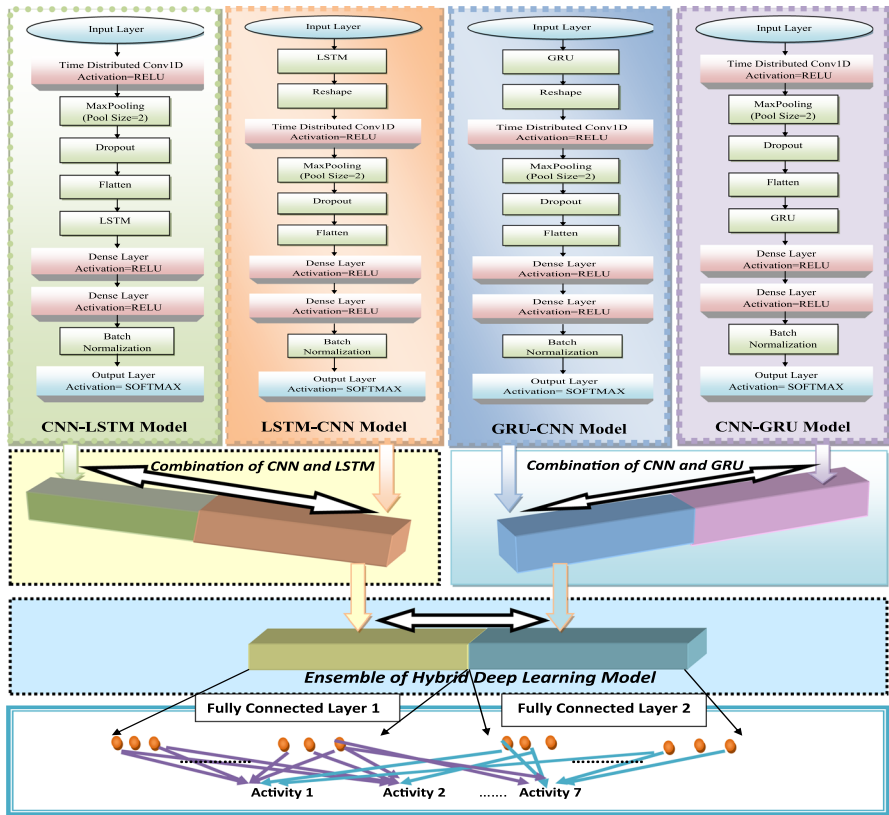


Fig. 5 Proposed hybrid deep learning architecture

Table 3 Set of parameters for CNN

Parameters	Value
α	0.001
β_1	0.9
β_2	0.999
Epsilon	10^{-6}
Epoch	170
Batch size	258

4.5.2 Precision

Precision is the ratio of the predicted positive observation to the total predicted positive observation, it is defined as follows:

Table 4 Set of parameters for CNN-LSTM

Parameters	Value
α	0.001
β_1	0.8
β_2	0.899
Epsilon	10^{-7}
Epoch	175
Batch size	264

Table 5 Set of parameters for CNN-GRU

Parameters	Value
α	0.001
β_1	0.9
β_2	0.999
Epsilon	10^{-8}
Epoch	175
Batch size	260

Table 6 Set of parameters for GRU-LSTM

Parameters	Value
α	0.001
β_1	0.9
β_2	0.999
Epsilon	10^{-7}
Epoch	175
Batch size	2564

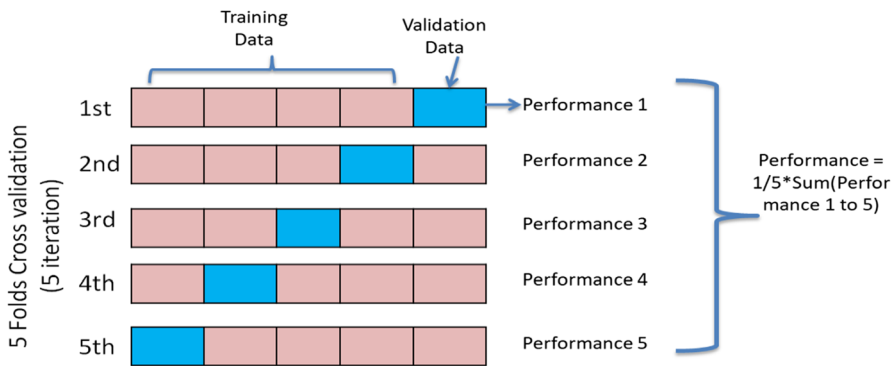


Fig. 6 Fivefold cross-validation Architecture

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (15)$$

4.5.3 Recall

Recall is the ratio of the true positive observation to the total actual positive observation, it is defined as follows:

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (16)$$

4.5.4 F1 score

F1 score on the other hand is the weighted average of the precision and recall. The formula for f-measure is calculated as:

$$\text{f-measure} = \frac{((1 + \beta^2) * \text{recall} * \text{precision})}{((\beta^2) * \text{recall} * \text{precision})} \quad (17)$$

where, the value assigned to β is 1.

4.5.5 Proposed algorithm

Algorithm 1 present details of proposed work.

Algorithm 1: proposed Hybrid Ensemble learning based 7 human walking activity recognition system

Result: Different walking activities reconnecting in term of accuracy, precision, recall and F score

Input: Raw data collected through IMU sensor on chest , left and right thigh.;

Output: Walking activities classification accuracy ;

Procedure;

Step 1 Pre-processed the data using extended kalman filter ;

Step 2 Prepare the data set in training , testing and validation set. ;

Step 3 Design of Hybrid deep learning model;

Step 4 Combined the performance using ensemble learning classifier;

Step 5 Finally calculate the performance metric;

5 Experiment results and discussions

The results were obtained on computer system with i8 8700U processor and 12.0 GB RAM. In first step, the collected data of 50 subject of 7 walking activities are validated.

5.1 Input tensor size

In next step, the tensor input is prepared for validation of deep learning model. Total 9 attribute of accelerometer reading around chest, left and right thigh is calculated for all 7 activities. The sampling frequency was 50 Hz so the input vector size was (50*9) and the total tensor size for 7 activities was (700*450).

5.2 Model fitting and classification

The classification of different human walking activities was performed using seven different deep learning classifiers namely, CNN, LSTM, GRU, CNN–LSTM, LSTM–CNN, CNN–GRU, GRU–CNN and achieved 97.26%, 90.67%, 77.38%, 97.83%, 94.35%, 97.64%, 96.98% accuracy, respectively, on our data set. Tables 3, 4, 5 and 6 shows the hyperparameters used by abovementioned four hybrid deep learning models. Later, the 3 stage ensemble learning-based classifier is designed which has reported testing accuracy 99.34%, highest among all classifier.

5.3 Performance evaluation metrics

The accuracy of all the classification techniques is computed to compare the performance of each of the algorithms implemented. The results also incorporated precision, recall, F1 score, and support for all the labeled outputs. Confusion matrix is also displayed for all the abovementioned classification techniques.

5.3.1 Performance measurement in terms of classifier accuracy

The data set consists of seven different walking activities. There are total 50 subjects data across all 7 classes. Figure 7 shows the classification accuracy report for different classifiers for individual activities. Table 7 represents the classification accuracy report for different classifiers on out data set and MHEALTH data set. The overall accuracy of the Ensemble learning classifier was found to be 99.34%. To improve the overall classification accuracy of different activity gait, the performance of all different classifier has been combined for different hyperparameter. The average of all five classifier has given overall 99.34% accuracy in 4 different trials. Figure 13 shows the individual walking activities accuracy of different classifiers.

We have provided the confusion matrices for all the classifiers for human walking activities classification. We have computed precision, accuracy, recall, $f - 1$ score and support value for each classifier and for each activity and comparison is made. Table 8 shows the detailed performance over various parameters. The model accuracy and loss plot with output as confusion matrix are plotted for all the four hybrid deep learning models, i.e., CNN–LSTM, LSTM–CNN, CNN–GRU and GRU–CNN and ensemble learning. In the case of CNN–LSTM, we inspect a steep drop in the values of loss of training and validation data with in first 3–4 epochs. A gradual increase in the accuracy of training and validation data is depicted in first 5 epochs

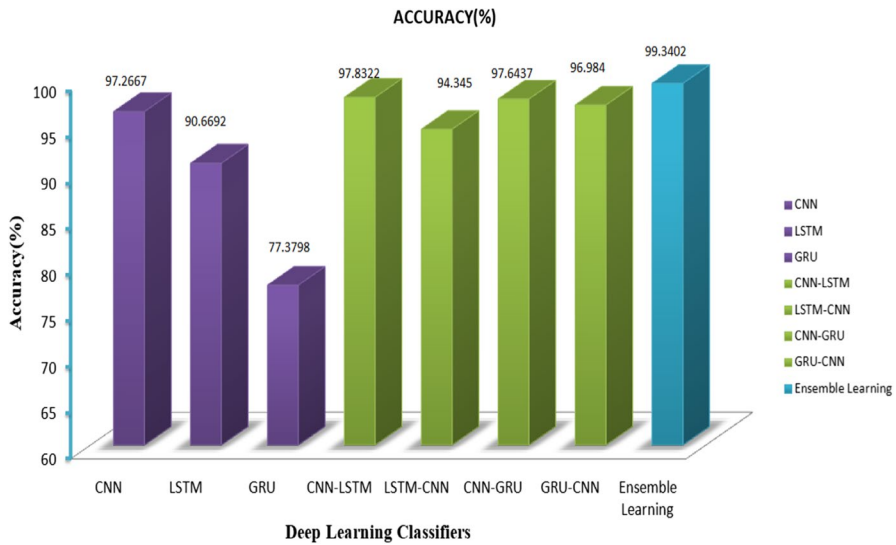


Fig. 7 Accuracy classification report

Table 7 Comparison with state of the art models

Data set	Methods	Accuracy (%)
Proposed HAG	CNN [21]	97.26
	LSTM [21]	90.67
	GRU [21]	77.38
	CNN-LSTM [21]	97.83
	LSTM-CNN [22]	94.35
	CNN-GUR [21]	97.64
	GRU-CNN [21]	96.98
	Proposed ensemble method	99.34
MHEALTH	CNN [21]	91.66
	LSTM [21]	86.89
	GRU [21]	81.77
	CNN-LSTM [21]	91.66
	LSTM-CNN [22]	89.0
	CNN-GUR [21]	91.66
	GRU-CNN [21]	92.53
	Proposed ensemble method	94

and then we see a constant curve. After that some oscillations are seen in loss of training and validation data and a very slight gain of accuracy is observed. Figure 8 shows the confusion matrix and loss plot for CNN-LSTM model.

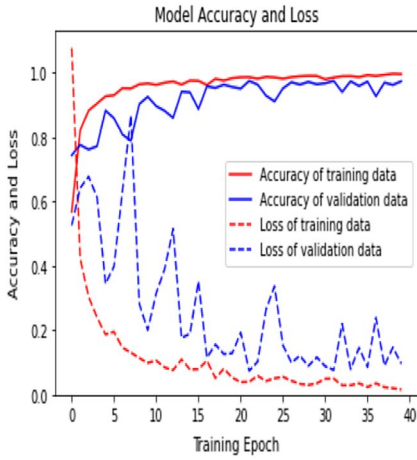
In the output of LSTM-CNN, till first 15 epochs a rapid decrease in the loss of training and validation data is seen, in the same time accuracy boosted up to 50%. Then, till first 60 epochs we can see an oscillatory decrease in the values of loss.

Table 8 Matrices showing individual results

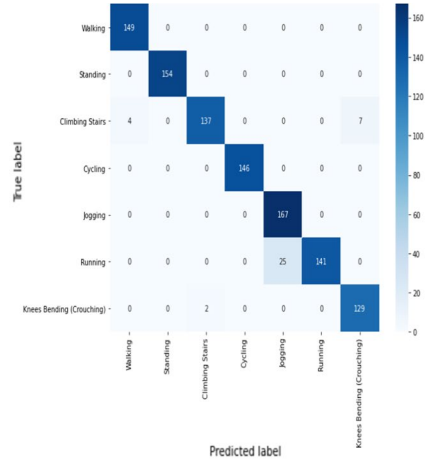
Model	Parameters	Walking	Standing	Climbing stairs	Cycling	Jogging	Running	Knees bending (crouching)
CNN	Precision	0.97	1.0	0.99	1.0	0.97	0.95	0.97
	Recall	1.0	1.0	0.91	1.0	0.95	0.99	0.99
	F1	0.98	1.0	0.94	1.0	0.96	0.97	0.97
	Support	158	158	143	159	149	147	147
LSTM	Precision	0.89	0.93	0.84	0.99	0.97	0.86	0.88
	Recall	1.0	1.0	0.76	0.99	0.84	0.99	0.77
	F1	0.94	0.97	0.80	0.99	0.90	0.92	0.82
	Support	154	157	152	161	151	144	142
GRU	Precision	0.94	0.99	0.87	0.99	0.98	0.81	0.92
	Recall	1.0	0.99	0.86	0.99	0.77	1.0	0.85
	F1	0.97	0.99	0.86	0.99	0.86	0.89	0.88
	Support	154	157	152	161	151	144	142
LSTM–CNN	Precision	0.99	0.95	0.97	1.0	0.94	0.99	0.90
	Recall	0.98	0.99	0.90	1.0	0.99	0.94	0.93
	F1	0.99	0.97	0.94	1.0	0.97	0.97	0.91
	Support	154	157	152	161	151	144	142
CNN–LSTM	Precision	0.98	0.99	0.97	1.0	0.98	0.97	0.99
	Recall	1.0	1.00	0.96	1.0	0.97	0.99	0.97
	F1	0.99	0.99	0.97	1.0	0.98	0.98	0.98
	Support	158	158	143	159	149	147	147
CNN–GRU	Precision	0.87	0.99	0.93	0.99	0.99	0.99	1.0
	Recall	1.0	1.0	0.85	1.0	0.99	0.99	0.90
	F1	0.93	0.99	0.89	0.99	0.99	0.99	0.95
	Support	154	157	152	161	151	144	142
GRU–CNN	Precision	0.99	0.98	1.0	0.99	0.99	0.98	0.89
	Recall	1.0	0.99	0.88	1.0	0.98	1.0	0.98
	F1	0.99	0.99	0.93	0.99	0.99	0.99	0.93
	Support	154	157	152	161	151	144	142
Ensemble learning	Precision	1.00	0.99	0.99	1.00	0.99	0.98	0.99
	Recall	1.0	0.99	1.00	1.00	0.99	0.99	0.98
	F1	1.00	0.99	0.99	1.00	0.99	0.99	0.99
	Support	150	162	157	148	170	126	148

Accuracy meanwhile takes a constant curve. Figure 9 shows the confusion matrix and accuracy and loss curve using LSTM–CNN model.

Output of CNN–GRU takes almost similar pattern as GRU–CNN, while this model gives us 100% accuracy. We can see the confusion matrix Fig. 10 and the characteristic plot Fig. 11. The ensemble learning model is best suited for the activity recognition problem. Figure 12 shows the best output for ensemble learning model. When evaluating the classification models, it is required to test their accuracy

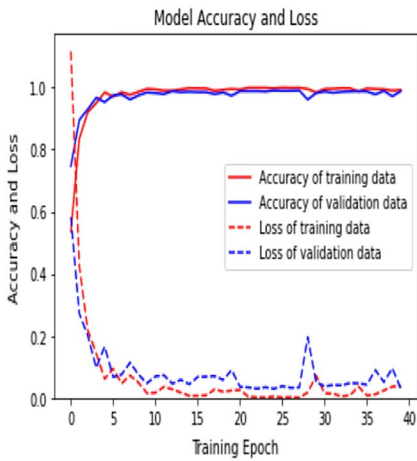


(a) Accuracy and loss curve for CNN-LSTM model.

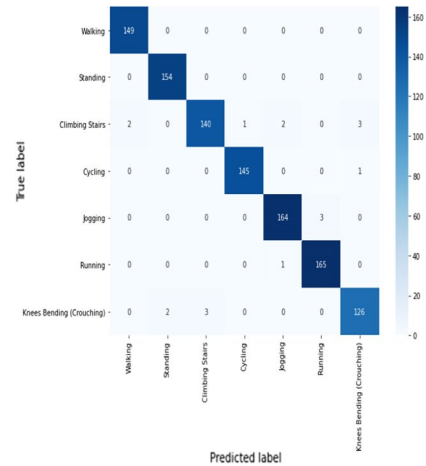


(b) Confusion matrix for CNN-LSTM model.

Fig. 8 CNN-LSTM model



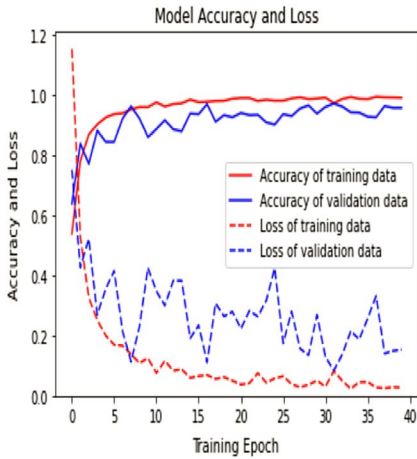
(a) Accuracy and loss curve for LSTM-CNN model.



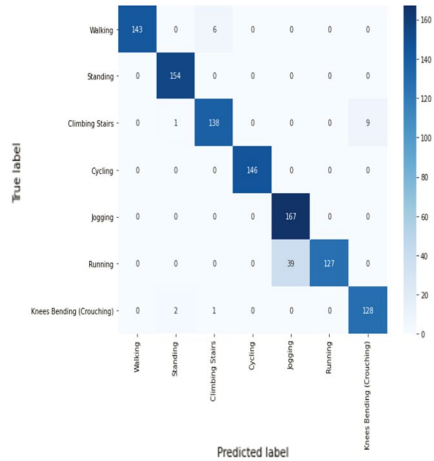
(b) Confusion matrix for LSTM-CNN model.

Fig. 9 LSTM-CNN model

on different data than that was used initially to train the model. This confirms that our model is not “overfit” to one small set of data. The proposed data set is validated with Mhealth data set. Table 7 shows the different model accuracy comparison of our data set with MHEALTH data. It shows these method performance better for our data set. The results also show the superiority of ensemble learning-based classifier (Fig. 13).



(a) Accuracy and loss curve for CNN-GRU model.

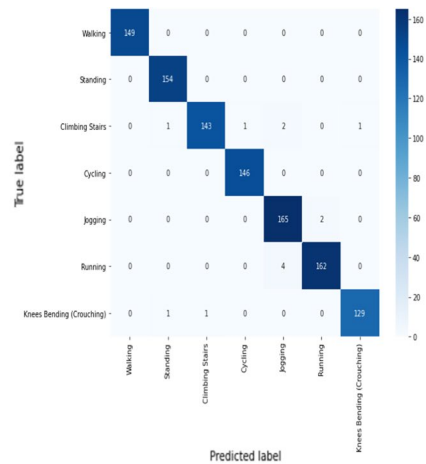


(b) Confusion matrix for CNN-GRU model.

Fig. 10 CNN-GRU model



(a) Accuracy and loss curve for GRU-CNN model.

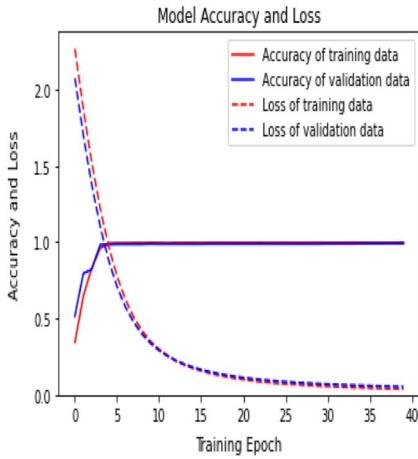


(b) Confusion matrix for GRU-CNN model.

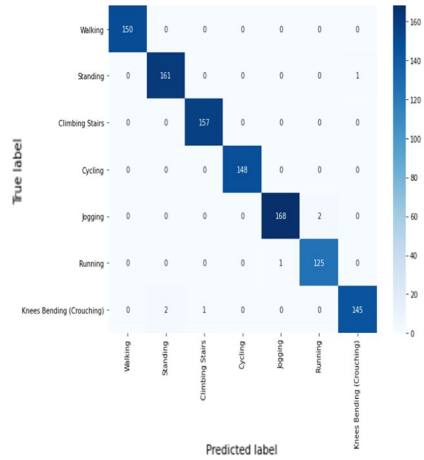
Fig. 11 GRU-CNN model

6 Conclusion

This paper has considered 50 subjects data for seven different walking activities. The data were collected through IMU sensor, placed at chest, left thigh, and right thigh. Total 100 samples are collected for all 11 activities. The sampling rate considered was 50 Hz. Following 7 walking activities were performed for



(a) Accuracy and loss curve for Ensemble Learning model.



(b) Confusion matrix for Ensemble Learning model.

Fig. 12 Ensemble learning model

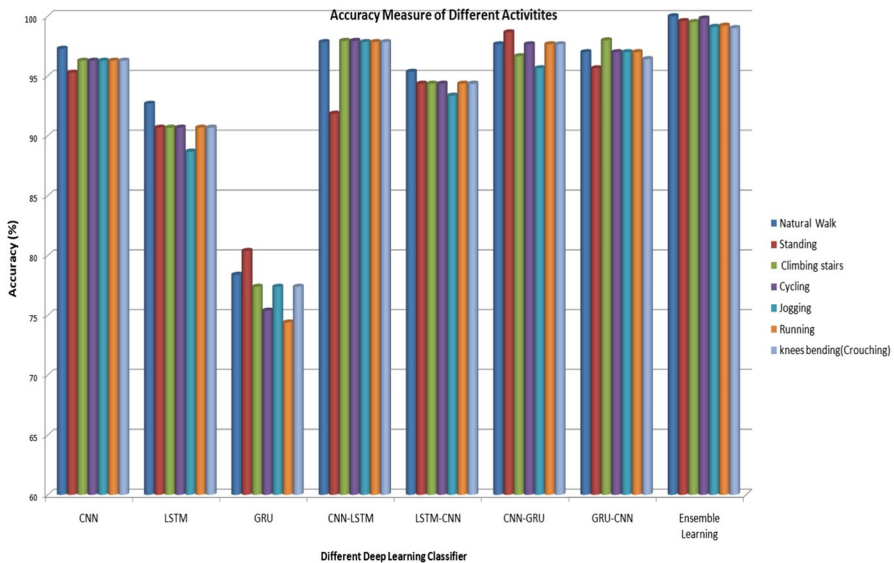


Fig. 13 Individual activity accuracy classification report

all the 50 subjects:.. These activities are (i) natural walk, (ii) standing, (iii) climbing stairs, (iv) cycling, (v) jogging, (vi) running, (vii) knees bending(Crouching). The data were classified using four Hybrid deep learning models namely, CNN-LSTM, CNN-GRU, LSTM-GRU, LSTM-CNN, and ensemble of all model. The accuracy was achieved for different deep learning models are CNN, LSTM,

GRU, CNN–LSTM, LSTM–CNN, CNN–GRU, GRU–CNN have achieved 97.26%, 90.67%, 77.38%, 97.83%, 94.35%, 97.64%, 96.98% accuracy, respectively, on our data set. The paper has also implemented the ensemble learning by providing the combined average results of different classifier and providing the dropout. The proposed ensemble learning-based hybrid deep learning framework has provided a promising classification accuracy of 99.34% over other models. It provides the generic solution and removes the dependency on more on hyperparameter and data dependence. The ensemble technique has reduced the variance and model complexity.

6.1 Future scope

The work further can be extended for stable robot walk generation, automatic person tracking, clinically person health monitoring, surveillance of person based on gait, restoring of elderly and crouch walk and pedestrian navigation, etc. As a future scope, the economical reliable and sustainable solution can be developed by imparting the human behavior and learning using hybrid deep learning algorithm for automatically gait activities recognition. Gait-related walking activities are very important for analysis of postural instability, repairment of gait abnormality, diagnosis of cognitive declination, and enhance the cognitive ability of human-centered humanoid robot system. Though this proposed system is a prototype, but the preliminary result is encouraging and leads to development of an integrated system for tracking automatic gait activities. The system may consist of IMU sensor with processing, display device and in a software part there will be a pre-recorded perfect walking activity recognition so that subject may follow the rehabilitation activities perfectly.

Acknowledgements The author(s) would like to thank all the participants who have allowed us to capture the data using a wearable device. The author(s) would like to express thanks to Human locomotion analysis laboratory of the Institute of Technology Gopeshwar, Uttarakhand and Human motion capturing and analysis unit of MANIT Bhopal for providing the opportunity to collect data and providing the computing facility.

Funding This paper is the results of Project Funded by SERB, DST Govt. of India under the schema of Early career award to Dr. Vijay Bhaskar Semwal, MANIT Bhopal, DST No: ECR/2018/000203 dated on 04/06/2019.

Declaration

Conflict of interest The authors declare there is no conflict regarding this research paper.

References

1. Ahmed MH, Sabir AT (2017) Human gender classification based on gait features using kinect sensor. In: 2017 3rd IEEE International Conference on Cybernetics (Cybconf). IEEE, pp 1–5
2. Semwal VB, Raj M, Nandi GC (2015) Biometric gait identification based on a multilayer perceptron. *Robot Auton Syst* 65:65–75
3. Semwal V. B (2017) Data driven computational model for bipedal walking and push recovery. [arXiv:1710.06548](https://arxiv.org/abs/1710.06548)

4. Semwal VB, Katiyar SA, Chakraborty R, Nandi GC (2015) Biologically-inspired push recovery capable bipedal locomotion modeling through hybrid automata. *Robot Auton Syst* 70:181–190
5. Semwal VB, Bhushan A, Nandi G (2013) Study of humanoid push recovery based on experiments. In: 2013 International Conference on Control, Automation, Robotics and Embedded Systems (CARE). IEEE, pp 1–6
6. Guo Y, Wu X, Shen L, Zhang Z, Zhang Y (2019) Method of gait disorders in Parkinson's disease classification based on machine learning algorithms. In: 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). IEEE, pp 768–772
7. Patil P, Kumar KS, Gaud N, Semwal VB (2019) Clinical human gait classification: extreme learning machine approach. In: 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT). IEEE, pp 1–6
8. Semwal VB, Nandi GC (2016) Generation of joint trajectories using hybrid automate-based model: a rocking block-based approach. *IEEE Sens J* 16(14):5805–5816
9. Nandi GC, Semwal VB, Raj M, Jindal A (2016) Modeling bipedal locomotion trajectories using hybrid automata. In: 2016 IEEE Region 10 Conference (TENCON). IEEE, pp 1013–1018
10. Li X, Yuan Z, Zhao J, Du B, Liao X, Humar I (2021) Edge-learning-enabled realistic touch and stable communication for remote haptic display. *IEEE Netw* 35(1):141–147
11. Gupta JP, Polytool D, Singh N, Semwal VB (2014) Analysis of gait pattern to recognize the human activities. *IJIMAI* 2(7):7–16
12. Semwal VB, Nandi GC (2015) Toward developing a computational model for bipedal push recovery—a brief. *IEEE Sens J* 15(4):2021–2022
13. Hsu W-C, Sugiarto T, Lin Y-J, Yang F-C, Lin Z-Y, Sun C-T, Hsu C-L, Chou K-N (2018) Multiple-wearable-sensor-based gait classification and analysis in patients with neurological disorders. *Sensors* 18(10):3397
14. Mekruksavanich S, Jitpattanakul A, Youplao P, Yupapin P (2020) Enhanced hand-oriented activity recognition based on smartwatch sensor data using LSTMs. *Symmetry* 12(9):1570
15. Kwapisz J, Weiss G, Moore S (2010) Activity recognition using cell phone accelerometers. *SigKDD Explor Newsllett* 12(101145):1964897–1964918
16. Papavasileiou I, Zhang W, Wang X, Bi J, Zhang L, Han S (2017) Classification of neurological gait disorders using multi-task feature learning. In: 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems And Engineering Technologies (CHASE). IEEE, pp 195–204
17. Semwal VB, Gaud N, Nandi G (2019) Human gait state prediction using cellular automata and classification using ELM. In: *Machine Intelligence and Signal Analysis*. Springer, pp 135–145
18. Semwal VB, Singha J, Sharma PK, Chauhan A, Behera B (2017) An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification. *Multimed Tools Appl* 76(22):24457–24475
19. Semwal VB, Kumar C, Mishra PK, Nandi GC (2016) Design of vector field for different sub-phases of gait and regeneration of gait pattern. *IEEE Trans Autom Sci Eng* 15(1):104–110
20. Chen Z, Li G, Fioranelli F, Griffiths H (2018) Personnel recognition and gait classification based on multistatic micro-Doppler signatures using deep convolutional neural networks. *IEEE Geosci Remote Sens Lett* 15(5):669–673
21. Semwal VB, Mondal K, Nandi GC (2017) Robust and accurate feature selection for humanoid push recovery and classification: deep learning approach. *Neural Comput Appl* 28(3):565–574
22. Poschadel N, Moghaddamia S, Alcaraz JC, Steinbach M, Peissig J (2017) A dictionary learning based approach for gait classification. In: 2017 22nd International Conference on Digital Signal Processing (DSP). IEEE, pp 1–4
23. Semwal VB, Chakraborty P, Nandi GC (2015) Less computationally intensive fuzzy logic (type-1)-based controller for humanoid push recovery. *Robot Auton Syst* 63:122–135
24. Wang X, Zhang J, Yan WQ (2019) Gait recognition using multichannel convolution neural networks. *Neural Comput Appl* 32:14275–14285
25. V B, Gupta V, Semwal VB (2021) Wearable sensor based pattern mining for human activity recognition: deep learning approach. *Ind Robot* 48(1)
26. Gupta A, Semwal VB (2020) Multiple task human gait analysis and identification: ensemble learning approach. In: *Emotion and information processing*. Springer, pp 185–197
27. Wang X, Yan K (2020) Gait classification through CNN-based ensemble learning. *Multimed Tools Appl* 80:1565–1581

28. Sun L, Yuan Y-X, Zhang Q, Wu Y-C (2018) Human gait classification using micro-motion and ensemble learning. In: IGARSS 2018–2018 IEEE International Geoscience And Remote Sensing Symposium. IEEE, pp 6971–6974
29. Wang X, Yan WQ (2020) Cross-view gait recognition through ensemble learning. *Neural Comput Appl* 32(11):7275–7287
30. Shu J, Hamano F, Angus J (2014) Application of extended Kalman filter for improving the accuracy and smoothness of Kinect skeleton-joint estimates. *J Eng Math* 88(1):161–175
31. Banos O, Villalonga C, Garcia R, Saez A, Damas M, Holgado-Terriza JA, Lee S, Pomares H, Rojas I (2015) Design, implementation and validation of a novel open framework for agile development of mobile health applications. *Biomed Eng Online* 14(2):1–20
32. Banos O, Garcia R, Holgado-Terriza JA, Damas M, Pomares H, Rojas I, Saez A, Villalonga C (2014) Mhealthdroid: a novel framework for agile development of mobile health applications. In: *International Workshop on Ambient Assisted Living*. Springer, pp 91–98

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Vijay Bhaskar Semwal¹  · Anjali Gupta¹  · Praveen Lalwani²

✉ Vijay Bhaskar Semwal
vsemwal@manit.ac.in

Anjali Gupta
anjalig376@gmail.com

¹ Maulana Azad National Institute of Technology, Bhopal, Madhya Pradesh, India

² VIT, Bhopal, Madhya Pradesh, India