Smartphone Sensor-based Human Activity Recognition System using BRNN-LSTM Method

Dr. P. Maragathavalli[1] and Ms. R.Vivitha[2]

[1] *Assistant Professor, Department of Information Technology Engineering, Pondicherry Engineering College, Kalapet, Pondicherry*

[2] *Student Member, Department of Information Technology Engineering, Pondicherry Engineering College, Kalapet, Pondicherry*

Abstract: Smartphones are quickly becoming the most important communication device in people's lives in today's world. Human Activity Recognition has grown in popularity as a field of study in a variety of areas, including medical care, tracking, and education. The sensors in smartphones allow us to use them for a wide range of applications. Healthcare is one of the major domains where human activity recognition is widely used. In this paper, a human activity recognition system has been developed that can detect six activities of daily living (ADL) along with human fall. Human fall occurs due to an accident that can cause serious injuries which may lead to significant medical problems when the issue is not addressed properly. The proposed system uses a variant of deep learning technique to detect human activities and human fall. The accuracy is significantly increased by nearly 4% when compared with previous results.

Keywords: Activities of Daily Life; Smartphone sensors; Accelerometer; Gyroscope; Movement monitoring ; Healthcare; Bidirectional LSTM

1. Introduction

In recent times, smartphones have become an indispensable part in the life of humans. The daily routine of an individual can be captured by using smartphones. The accelerometer and gyroscope are the most widely used sensors. The in-built sensors in smartphones enable us to track the individual's movement. This information can be utilised to recognise several day-to-day routine activities like standing, walking, laying, sitting, walking upstairs and walking downstairs[2]. The same information is also utilized for detecting human fall. Data from a person's daily life may be used to investigate their way of life and the sequence of their bodily activities.Since lack of physical activity causes the majority of chronic diseases, tracking everyday movements will quickly identify any deviations from normal activities. Identifying human falls will also assist in providing appropriate assistance to those who are in need. HAR has numerous applications in the various fields such as elderly care, monitoring the patient, education, military, and rehabilitations[2][3][8].

Fall detection systems[3] are systems that help in monitoring human activity and immediately alert at a time of catastrophe. The key objectives of fall detection systems are to distinguish between activities of daily living (ADL) and fall events (forewarning during their occurrence). These systems aid to send immediate notification to medical entities, caretakers, family members to ensure that timely care can be given.

This paper proposes an enhanced human activity recognition model for fall detection using modified Bidirectional Recurrent Neural Network (BRNN) incorporated with Long Short-Term Memory (LSTM), a Deep Learning (DL) approach. First of all the data for the proposed system is collected from the smartphone sensors such as the accelerometer and gyroscope. Secondly, the data is then preprocessed and segmented. Then the system is trained and tested for six human activities including fall. The Mobifall dataset which contains sensor data of 66 subjects performing four different falls and twelve different ADLs is used. Finally, metrics like accuracy, f-measure, precision and confusion matrix are used to measure the performance of the system[2][3][8].

The organization of the paper is as follows: Section 2 presents existing and a few of the correlated works. Section 3 provides the outline of the modified BRNN-LSTM system proposal. Experimentation and evaluation of the model is presented in Section 4. Section 5 bestows the outcome. Section 6 presents the conclusion with works for the future direction.

2. Background and Related Work

The basic steps involved in human activity recognition system comprises the following stages/steps: gathering data, pre-processing data, feature extraction, categorizing activities and interpretation of results, as shown in Figure 1.

Variety of sensors are used to gather the input data. The next step is pre-processing which helps in elimination of noise and dealing with erroneous values which can be done by applying data cleansing methods. Cleansed data is then segmented into windows. Segmentation of windows is done using many ways like sliding windows, event based windows and energy windows. Following the completion of pre-processing and segmentation, feature extraction is performed to improve the performance of classification algorithms. To accomplish this task, deep learning methods are used [2][3]. The features extracted are used to classify the activities using various classification algorithms and metrics such as accuracy, F1-score, precision and confusion matrix are used to assess the system's performance [1][2][3].

Fig. 1. Steps involved in Human Activity Recognition

The ability to perform automated feature extraction is one of the main reasons for the popularity of deep learning methods in recent times. In the existing works, Human Activity Recognition (HAR) is done by various Deep Learning (DL) Techniques namely CNN[1][3], RNN, RNN-LSTM[2][4] as shown in Table 1. The main issue in the existing work is the issue of high computation power that makes real-time predictions arduous as most systems are incapable to handle it.

Human Activity Recognition is a time series problem such that the sequence of sensor readings are in time T. LSTM[2] cells can catch relationships in time dependent data without having to mix the timesteps together as a 1D convolutional neural network (CNN)[3] would do. At present, as "big data'' emerges, LSTM architecture can offer better outcomes and in various unrealized applications. Though researchers have made significant progress in the field of HAR, in recent times, there is still hope for potential improvements. The present state is predicted by the LSTM cells from the information of the former state. Apparently, we can observe that in real-world human trajectories are ceaseless. As a result, memory cells running in a single direction will not be able to acquire the necessary information. The proposed framework employs a novel approach that combines bidirectional recurrent neural networks with long short-term memory cells to ensure the uprightness of the data fed to the various layers of neural networks, while also improving the accuracy of the existing model.

Table 1. Comparison of existing deep learning techniques

Sl.No.	NAME OF THE JOURNAL, YEAR	PAPER TITLE	TECHNIQUES USED	METRICS	MERITS AND DEMERITS
1.	Computing, Springer 2021	Multi-input CNN-GRU based human activity recognition using wearable sensors	Convolutional Neural Network and Gated Recurrent Unit	Accuracy F1-score	- Provides accurate results. - Low learning efficiency.
2.	Mobile Network Application, Springer 2020	Deep Learning Models for Real-time Human Activity Recognition with Smartphones.	Mobile Edge Computing and Convolutional Neural Network	Accuracy F1-score	- Provides real-time supervision. - Sensor used has a fixed range and time constant.
3.	IEEE Access, IEEE 2020	On the Personalization of Classification for Human Activity Recognition	Personalization models using Adaboost classifier	Accuracy	- Experimented on a new personalisation method. - Used only one sensor
4.	XXVI Brazilian Congress on Biomedical Engineering, Springer 2019	Human Activity Recognition based on Convolutional Neural Network	Convolutional Neural Network	Accuracy, precision	- Lower computational cost. - Information loss.
5.	International Conference on Computational Intelligence and Data Science (ICCIDS), Elsevier 2019	A Lightweight Deep Learning Model for Human Activity Recognition on Edge Devices	Deep learning approach using Recurrent Neural Network and Long Short Term Memory	Accuracy, precision, recall, F1-score	- Reduces communication latency, cost and network traffic. - Does not support multi-sensor data.
6.	Robotics and Automation Letters, IEEE 2019	HMFP-DBRNN: Real-Time Hand Motion Filtering and Prediction via Deep Bidirectional RNN	Deep bidirectional Recurrent Neural Network approach	Accuracy F1-score	- Produces precise outputs. - Overfitting of data occurs.
7.	Sensors Journal, IEEE Access 2020	Advanced Sensing and Human Activity Recognition in Early Intervention and Rehabilitation of Elderly People.	Supervised machine learning approach	Accuracy	- Provides accurate results. - More sensors are required.

3. Illustration of the Proposed System

The proposed system is evolved using BRNN and LSTM. It has two hidden layers and 30 neurons. Each module in the system is explained as follows.

3.1. Recurrent Neural Network (RNN)

RNN algorithm is used to solve problems involving continuous data. As HAR is a time series problem, the input is in sequential form. It comprises three layers - i/p, hidden, and o/p, with hidden layers containing several nodes. Figure 2 shows the visual representation of a RNN node. The present hidden state value h_t and the output value o_t is generated by the generating function function of each hidden node t is calculated using formula:

$$h_t = \varepsilon(w_{hh}h_{t-1} + w_{ih}x_t + b_h) \quad (1)$$

$$y_t = \varepsilon(w_{ho}h_t + b_o) \quad (2)$$

where, x_t = new input state, h_{t-1} = old hidden state, h_t = new hidden value,

w_{hh} =hidden-hidden state weight, w_{ih} = input-hidden state weight, w_{ho} = hidden-output state weight, b_h = hidden state bias, b_o = output state bias, o_t = new output state, ε = activation function. ε is called an activation function.

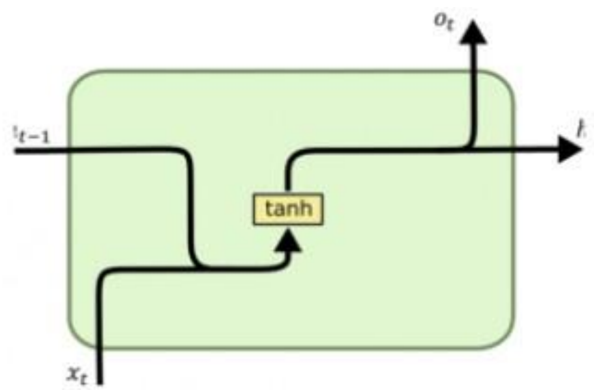

Fig. 2. Sample Node of Recurrent Neural Network (RNN)

3.2. Long Short Term Memory (LSTM)

The main drawback of using RNN is it's inability to handle exploding and vanishing gradients which decreases the performance of the network model in wide-range of temporal dependencies between inputs and human activities. Use of LSTM enables us to overcome the drawback of RNN, by using LSTM memory cells instead of RNN nodes. LSTM memory cells are shown in Figure 3. Each memory cell's characteristics are administered by the gates in LSTM cells. The activation functions of gates control the state of each cell. Input gate, forget gate and output gates are the gates to which the input values are fed. Activation function is represented by the symbol ε.

The following equations show each timestep(t) calculation of hidden state value.

$$i_t = \varepsilon_i(w_{xi}x_t + w_{hi}h_{t-1} + w_{ci}c_{t-1} + b_i) \quad (3)$$

$$f_t = \varepsilon_f(w_{xf}x_t + w_{hf}h_{t-1} + w_{cf}c_{t-1} + b_f) \quad (4)$$

$$c_t = f_tc_{t-1} + i_t\varepsilon_c(w_{xc}x_t + w_{hc}h_{t-1} + b_c) \quad (5)$$

$$o_t = \varepsilon_o(w_{xo}x_t + w_{ho}h_{t-1} + w_{co}c_t + b_o) \quad (6)$$

$$h_t = o_t\varepsilon_h(c_t) \quad (7)$$

$W_{xi}, w_{hi}, w_{ci}, w_{xf}, w_{hf}, w_{cf}, w_{xc}, w_{hc}, w_{xo}$, and w_{co}, are weights (w_{xi} = input to input weight, w_{hi} = hidden input weight), b_i, b_o, b_c, and b_f bias weights.

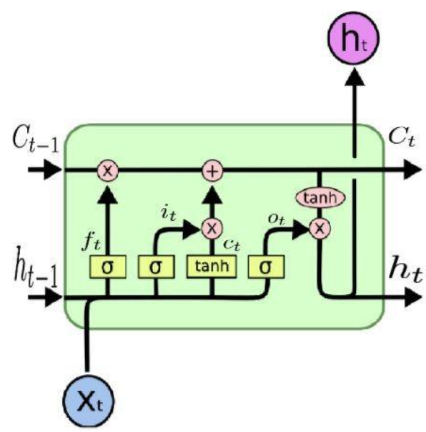

Fig. 3. Long-Short Term Memory (LSTM) Cell

3.3. Proposed modified BRNN-LSTM Model

Accelerometer and gyroscope readings are splitted into fixed size window T with 128 timesteps. The neural network input consists of reading (y_1, y_2, y_3,........,y_{T-1}, y_T) obtained at time T, where y_T is the reading obtained at any time instance t. The readings from the windows that are segmented are then given to the modified BRNN-LSTM network. The output is combined to single final output from different states using Rectified Linear Unit (ReLU). The model is compiled using categorical crossentropy and adam optimiser is used.

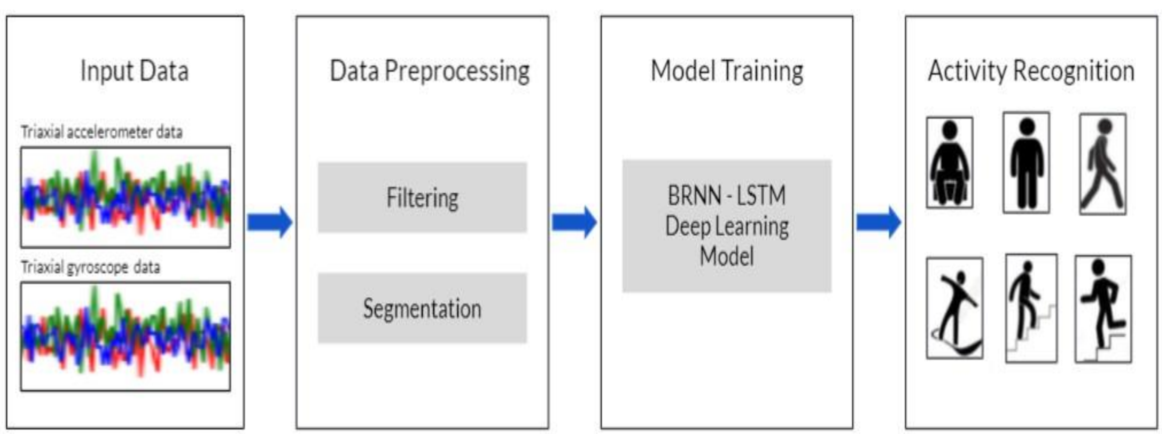

Fig. 4. Proposed System Design

As shown in Figure 5, improvement in the proposed method shows that not only the previous information but also the upcoming information is taken into account to achieve the current result. The fundamental approach of BRNN is to separate the neurons of a traditional RNN into two directions, one for forward states(positive time direction), and the other one for backward states (negative time direction). This helps in accurate prediction of human activities especially sitting and standing as those two activities have less accuracy in the existing methods.

```
model = keras.Sequential()
model.add(
    keras.layers.Bidirectional(
      keras.layers.LSTM(
          units=128,
          input_shape=[X_train.shape[1], X_train.shape[2]]
      )
    )
)
model.add(keras.layers.Dropout(rate=0.5))
model.add(keras.layers.Dense(units=128, activation='relu'))
model.add(keras.layers.Dense(y_train.shape[1], activation='softmax'))
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['acc'])
```

Fig. 5. Pseudo code for the modified BRNN-LSTM method

4. Experiment and Evaluation

The enhanced human activity recognition system for fall detection is executed using python 3.5 and tensorflow 1.7.

4.1. Dataset Description

Data from the accelerometer and gyroscope sensors of a smartphone were recorded. Mobifall is an open-source dataset that contains data gathered from a smartphone as participants engage in a variety of activities and fall. The data set comprises four types of falls, twelve distinct ADLs obtained from a total of 66 subjects. The tabulated description of the activities recorded are shown in Figure 6 shown below. The ADLs are divided into three sections such as mentioned in Figure 6.

Data set is taken from the URL https://bmi.hmu.gr/the-mobifall-and-mobiact-datasets-2.

Dataset	Simple Movements	Standard Normal Life Movements	Sporting Activities	Falls
MobiFall & MobiAct	- Sitting on a chair - Stepping in a car - Stepping out of a car - Standing	- Normal walking - Going downstairs - Going upstairs	- Jogging - Jumping	- Forwards (use of hands to dampen fall) - Forwards (first impact on knees) - Sideward bending legs - Backward (while trying to sit down)

Fig. 6. ADLs and falls executed by the experimental subjects

Different events have different x, y, and z coordinates. The plot for walking shows maximum spikes in the z direction, while the plot for jogging shows maximum spikes in the x and y directions. The activation function describes a neuron's output in terms of its input. 90% of the MobiFall data is used for training in the BRNN-LSTM layer, with the remaining 10% being used for testing.

4.2. Performance Metrics

The performance of the system is analyzed using widely used metrics: accuracy, precision and f-measure are given below:

Table 2. Performance metrics

Metrics	Formula
Accuracy: Number of correct predictions over total number of predictions.	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$ where, TP = true positives, TN = true negatives, FP = false positives, FN = false negatives.
Precision: Number of actual true predictions over total true predictions.	$Precision = \frac{1}{C}\left(\sum_{C=1}^{C}\frac{tp_c}{tp_c+fp_c}\right)$ where, C = total no.of classes, tp_c = true positives for class c, fp_c = false positive for class c.
F-measure: Harmonic mean of precision and recall.	$f\text{-}measure = \sum_{C=1}^{C} 2\left(\frac{n_c}{N}\right) * \frac{precision_c * recall_c}{precision_c + recall_c}$ where, N = total no.of inputs, n_c = no.of inputs in class c, $precision_c$ = precision value for certain class c, $recall_c$ = recall value for certain class c.

5. Result Analysis

This section presents the evaluation results of the proposed system. The confusion matrix of the system is shown in figure 7. This matrix shows the relation between correctly and wrongly predicted activities. In the confusion matrix, TP (True Positive) represents the number of positive activity predictions that are correctly predicted whereas FP (False Positive) gives the value for number of positive activity predictions that are predicted as negative. Similarly, TN (True Negative) is the number of negative reviews correctly predicted and FN (False Negative) is the number of negative activity predictions predicted as positive. The proposed system achieved an overall accuracy of about 91% and good accuracy for activities like sitting, walking and fall. The metrics calculated for each activity is shown in table 3.

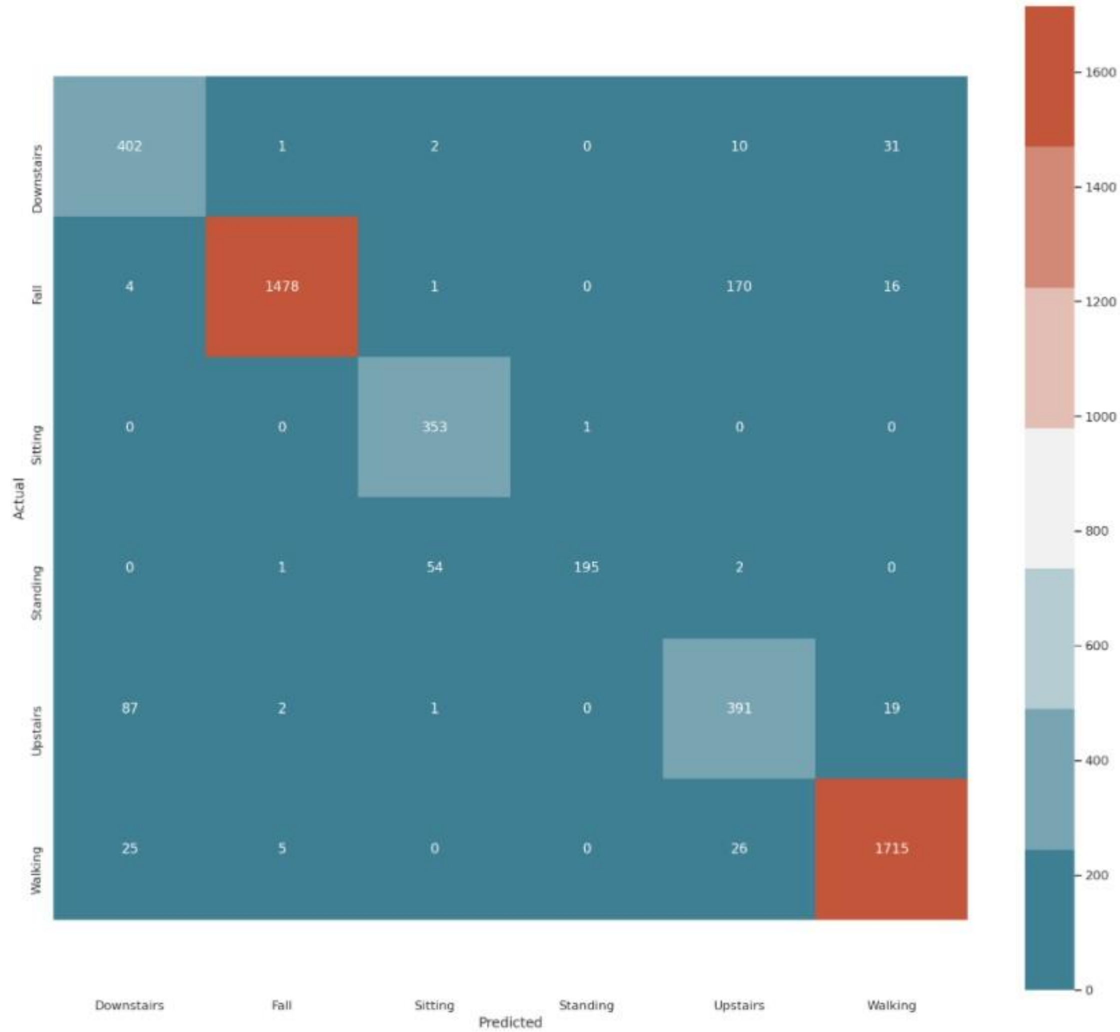

Fig. 7. Confusion Matrix

Table 3. Evaluation metrics calculated for each activity

Activity	Precision	Recall	F1-score
Downstairs	0.78	0.88	0.83
Fall	0.99	0.88	0.94
Sitting	0.86	0.99	0.92
Standing	0.99	0.76	0.86
Upstairs	0.66	0.78	0.71
Walking	0.97	0.97	0.97

The deep algorithm has a model loss graph drawn, which is a plot of loss on the training and validation datasets over training epochs. Note: The graph is drawn using code in python 3.x from the model executed and for validation 10% is for testing and 90% for training is given.

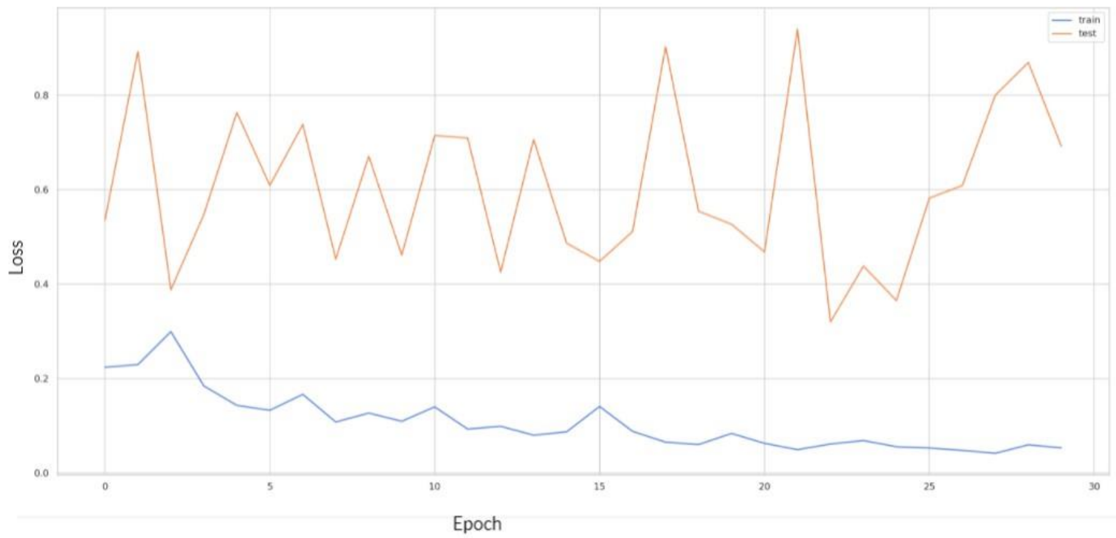

Fig. 8. Training progress vs Epochs

6. Conclusion

The HAR system using modified BRNN-LSTM for fall detection is developed in this paper. The system is able to predict five different activities along with fall successfully. The developed system produces nearly 4% more accuracy than many of the existing machine learning and deep learning methods by extracting the most representative features from the data.

Future directions can be included with more complex activities for recognition and also be extended to support multi-sensor data.

7. References

1. Dua, N., Singh, S.N. & Semwal, V.B. Multi-input CNN-GRU based human activity recognition using wearable sensors. Computing (2021). https://doi.org/10.1007/s00607-021-00928-8
2. Ferrari, D. Micucci, M. Mobilio and P. Napoletano, "On the Personalization of ClassificationModels for Human Activity Recognition," in IEEE Access, vol. 8, pp. 32066-32079, 2020, doi: 10.1109/ACCESS.2020.2973425.
3. Wan, S., Qi, L., Xu, X. et al. Deep Learning Models for Real-time Human Activity Recognition with Smartphones. Mobile Netw Appl 25, 743–755 (2020). https://doi.org/10.1007/s11036-019-
4. Coelho Y., Rangel L., dos Santos F., Frizera-Neto A., Bastos-Filho T. (2019), "Human Activity Recognition Based on Convolutional Neural Network". In: Costa-Felix R., Machado J., Alvarenga A(eds) XXVI Brazilian Congress on Biomedical Engineering. IFMBE Proceedings, Vol. 70/2. Springer, Singapore. https://doi.org/10.1007/978-981-13-2517-5_38
5. Preeti Agarwal, Mansaf Alam (2019), "A Lightweight Deep Learning Model for Human Activity Recognition on Edge Devices". International Conference on Computational Intelligence and Data Science (ICCIDS), Procedia Computer Science 167, pp.2364-2373, Elsevier, pp. 2364-2373.
6. Casilari E., Lora-Rivera R., García-Lagos F. (2019), "A Wearable Fall Detection System Using Deep Learning". In: Wotawa F., Friedrich G., Pill I., Koitz-Hristov R., Ali M. (eds) Advances and Trends in Artificial Intelligence. From Theory to Practice. IEA/AIE 2019. Lecture Notes in Computer Science, vol 11606. Springer, Cham. https://doi.org/10.1007/978-3-030-22999-3_39
7. Pienaar, S. W., & Malekian,R.(2019), "Human Activity Recognition using LSTM-RNN Deep Neural Network Architecture". arXiv preprint ArXiv: 1905:00599.
8. Ignatov, A. (2018), "Real-time human activity recognition from accelerometer data using Convolutional Neural Networks". Applied Soft Computing, 62, pp. 915-922.

9. P. Maragathavalli, R. Vivitha, M. Saranya and Uppu Dinesh, "Incorporating LSTM Method on Modified Deep Learning Technique for Sensor Based Human Activity Recognition System", International Conference on Frontiers in Engineering Science and Technology, ICFEST 2020, IOP Publishing, 1065 (2021) 012045, doi:10.1088/1757-899X/1065/1/012045.
10. Johan Wannenburg, et. al. (2018), "Wireless Capacitive-Based ECG Sensing for Feature Extraction and Mobile Health Monitoring", IEEE Sensors Journal, Vol. 18, No. 14, pp. 6023-6032.
11. S. Shahtalebi, S. F. Atashzar, R. V. Patel and A. Mohammadi (2019), "HMFP-DBRNN: Real-Time Hand Motion Filtering and Prediction via Deep Bidirectional RNN," in IEEE Robotics and Automation Letters, Vol. 4, No. 2, pp. 1061-1068. doi: 10.1109/LRA.2019.2894005.
12. Ahmed Nait Aicha, Gwenn Englebienne, Kimberley S.Van Schooten, Mirjam Pijnappels and Ben Krose (2019), "Deep Learning to Predict Falls in Older Adults Based on Daily-Life Trunk Accelerometry", in IEEE Sensors Journal, 18(5), 1654. doi: 10.3390/218051654
13. J. C. Dogan and M. S. Hossain (2019), "A Novel Two-Step Fall Detection Method Using Smartphone Sensors", 2019 IEEE International Conference on Smart Computing (SMARTCOMP), Washington DC, USA, pp. 434-438. doi: 10.1109/SMARTCOMP.2019.00083.
14. M. M. Hassan, A. Gumaei, G. Aloi, G. Fortino and M. Zhou (2019), "A Smartphone-Enabled Fall Detection Framework for Elderly People in Connected Home Healthcare," in IEEE Network, Vol. 33, No. 6, pp. 58-63. doi: 10.1109/MNET.001.1900100.
15. J. Lee and H. Tseng (2019), "Development of an Enhanced Threshold-Based Fall Detection System Using Smart phones With Built-In Accelerometers," IEEE Sensors Journal, vol. 19, no. 18, pp.8293-8302. doi: 10.1109/JSEN.2019.2918690.
16. Li,H., and Trocan, M. (2018), "Deep Learning of Smartphone Sensor Data for Personal Health Assistance", Microelectronics Journal.
17. "Mobifall Dataset" [Online]. Available: https://bmi.hmu.gr/the-mobifall-and-mobiact-datasets-2

Author Biography

Dr. P. Maragathavalli

She received her B.E degree in CSE from Bharathidasan University, M.Tech degree in CSE from Pondicherry University and PhD degree in CSE from Pondicherry University. She is working as Assistant Professor in the Department of Information Technology; Pondicherry Engineering College. She is a Life member of ISTE.

R. Vivitha

She is pursuing her B.Tech degree in the Department of Information Technology, Pondicherry Engineering College.