HUMAN ACTIVITY RECOGNITION USING HYBRID MODEL

Srikanta Swamy Pushpalatha (1), Shrishail Math (2)*

(1) Dept. of Information Science, Dr. Ambedkar Institute of Technology;
(2) Dept. of Computer Science and Engineering, Sri Krishna Institute of Technology
Bengaluru, India

* Corresponding Author: e-mail: pushpalat043@gmail.com

Abstract: The main task of the Human Activity Recognition (HAR) is to recognize the different actions performed by the individual. There are various machine learning and deep learning models which have been presented to predict the given activity of the human. The problem of the existing system is that they could not identify the activity when there is a sudden change in the activity. Hence, this paper proposes a model by using the deep learning concepts which can predict the activities and also can predict the sudden transition of one activity to another. This method uses the Convolution-Neural-Network Layer (CNN) and Gated-Recurrent-Unit (GRU) which identifies the changes in the activities collected by the sensors and gives us correct results. For the experimentation of this model, the University-of-California (UCI) Human Activity Recognition dataset has been used. This dataset comprises of various activities such as walking, sitting, standing etc. The results have compared with Human Activity Recognition System (HARS). The proposed model attained an accuracy of 96.79% whereas the HARS attained 95.99%. When compared with precision, recall and F1-score, the proposed model performed better than the existing model.

Key words: Monitoring, HAR, CNN, GRU, UCI

1. INTRODUCTION

The Human Activity Recognition has a vital role in inter-personal relations and interactions from one individual to another. The Human Activity Recognition gives the information about the personality, identity and psychological state of an individual and to extract this kind of information it is a very difficult task. The humans have an ability to recognize the various activities whereas the machines have to be trained and tested for each activity so that they can predict a correct outcome for any given instance. In the recent times, the study on the Human Activity Recognition has been increased due to its requirement in the human-computer interaction, healthcare, video surveillance and multiple activity recognition fields. The main aim of the Human Activity Recognition is to recognize the different activities performed by the individual from any image, video sequence or any dataset. The Human Activity Recognition model should also be able to classify the input data according to its correct category of activities. Human activities are classified into many types but we mainly focus on most common six activities. The most common human activity is walking, sitting, standing, running and these activities can be recognized easily. But a sudden instance where an individual starts running from a walking position the machines cannot predict easily. To resolve this issue, we have proposed a model, CNN-GRU which can identify the current activity and the sudden transition in the human activities. The significance of the proposed model is given below

- Propose a method using Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU) for identification of various actions and also identify the sudden change in the movement and predict the activity.
- Provide a good solution for classification and regression problem using the proposed CNN-GRU model.
- Provide a solution for time dependence when used in the noisy environment

2. LITERATURE SURVEY

The advancement of computer vision has led to the widespread adoption of HAR systems that utilizes wearable devices equipped with sensors as a standard part of daily life, for example, activity acknowledgment, controller, clinical checking and restoration exercises [5-8]. A variety of sensor-equipped wearable devices designed for use in everyday life and athletic competition are now in development. There is no need for additional sensors like cameras, radar, or infrared detectors because these sensors can monitor and recognize motion on their own. [9-11]. In [1], they have used Long-Short-Term-Memory (LSTM) for classifying the patterns in an unstructured text data. They have used the LSTM model to classify the feedbacks, complaints and recommendations of the customers. The results show that this model attained an accuracy of 91.67%. In [2], they have proposed an optimization algorithm called as MADEGA which dynamically generates neuro-fuzzy classification to detect the human status. This model was able to determine what is the status of the human using the optimized algorithm. This model outperformed all the existing data mining techniques. In [3], they have reviewed the various wireless sensor devices used for tracking pet and cattle animals. They have further classified them into various categories on the basis of domain, type and functionality. In this paper they have discussed more about how the wireless body area senor network work. In [4], they have proposed a method which will identify the posture of the person and asses the risks that might cause having that posture. This model identifies the posture using a body assessment tool, and Microsoft Kinect V2. The experimentation results show that this model can detect the postures of the student and reduce the risk of having any problems in their posture. The KEH can be useful in resolving the problem of excessive battery drain from wearable devices. With the help of the KEH, the battery consumption of wearable devices used in HAR can be cut by 79% [5]. Consequently, a great deal of studies has been directed for the most part on the cell phone sensors to get the information with respect to the human exercises. The dataset of the human exercises was arranged utilizing the cell phones [12]. A method based on histogram angles and Fourier structure has been proposed for extracting human behavior, and this method has been used to the UCI human action recognition data set. The SVM and KNN AI calculations were used to develop this technique for the task [13]. An approach has been described that uses the accelerometer in mobile devices to recognize motion. In this model, the information was split utilizing the accelerometer from the first information signal. The principal component analysis (PCA) is used in the model to restrict the factors and primarily divide the essential provisions of the various human activities in a specific period and recurrence for the setup of the six exercises. Seventy percent of the data was used for training with 6.11% accuracy, and thirty percent was used for testing with 92.10% accuracy, according to this method [14]. A technique wherein the UCI human cooperation acknowledgment knowledge base was utilized for the confirmation of their proposed model and for the arrangement of the dataset utilizing their model. There are as of now several examinations being done in the field of profound learning in HAR. Deep learning calculations can thereby extract the provisions from the dataset, but AI calculations require human intervention to do so. Thus, the information was extracted utilizing the sensors, the findings were reviewed and an effective model of HAR was constructed. We built a framework with the LSTM model to identify and categories the various physical activities in the dataset. The number of measurements was lowered applying the PCA approach and accomplished a precision of 97.64% [15]. With the use of the CNN model and the HAR dataset, researchers have developed a method to classify and rank the data with a 95.99% accuracy [16]. Together, the convolutional layer and the LSTM form a model with sophisticated neural organization. With a precision, recall and F1-score of 96.04 percent, 95.98 percent, and 96.01 percent, respectively [16], this model successfully deleted the items thereafter and organized the extracted provisions of the three datasets (OPPORTUNITY, UCI, and WISDM).

3. METHODOLOGY

In this section the methodology of the HAR model has been discussed. A basic model for the recognition of the human activities is given in the Figure 1. In this figure the basic human activities are accumulated using the sensors which is used as data for the detection of the activities. These human activities are then separated to the different motion data such as right knee motion data and left knee motion data. After this the pre-processing of the data is done in which the filtering and normalization and signal segmentation of the data accumulated by the sensor is done.

The next step is the feature extraction from the data and classify them using the deep learning model and predict which kind of human activity is being performed by the individual.

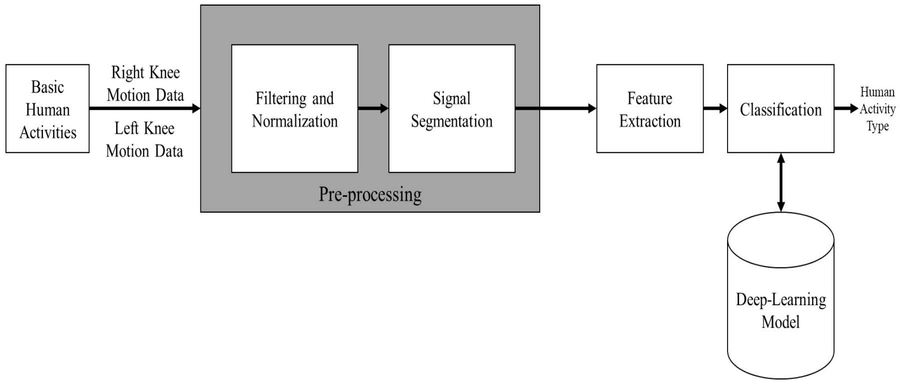

Figure 1. Basic model for recognition of human activity using machine learning model

For the HAR, there are two types. The first type is using the vision and the second type is using the sensor. The different activities can be identified using these two types. The vision based is done using the vision sensors or by a human individual who tells which activity is performed by the individual. Mostly the sensor-based gives accurate results because of the various sensos, wearable objects and dense sensing. The Figure 2. shows the recognition of human activity using vision and sensor-based.

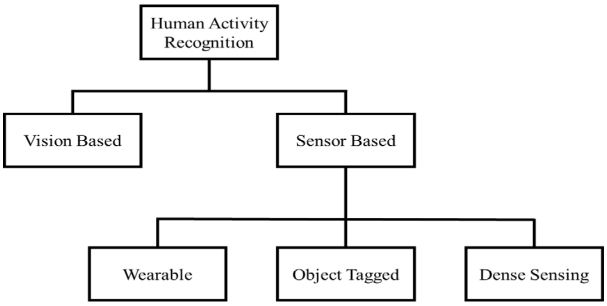

Figure 2. Recognition of Human Activity using vison and sensor based

The Figure 3 gives the architecture for our CNN-GRU model which is built on the deep learning framework. As in the architecture it can be seen that the K-sensor inputs are taken to classify and normalize the data before they are gone to the convolution layer. The K-sensor inputs are the six different activities such as walking, walking upstairs, walking downstairs, sitting, running, standing. In the convolution layer the K-sensor data is taken as input and all the feature are pre-processed before going to the gated recurrent unit. In the Convolution layer the features or the activities are identified using the algorithm and are classified. In the

GRU recurrent layer, the transition in the activities is identified and the output of the correct activity is given.

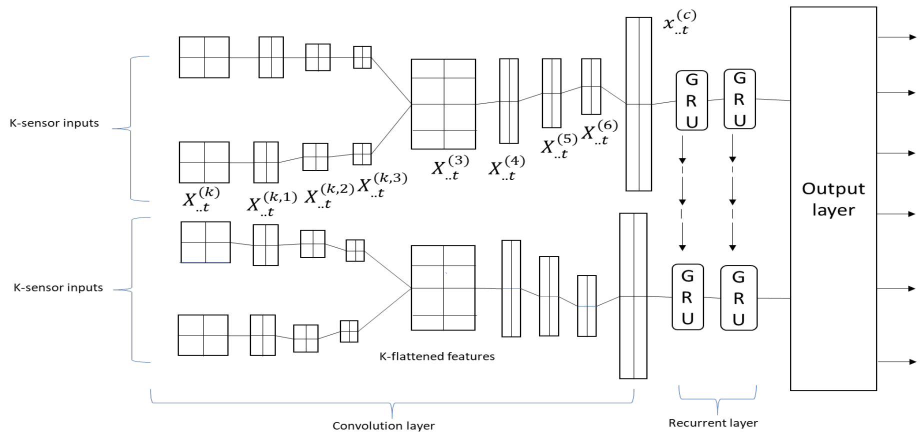

Figure 3. Architecture for our CNN-GRU model

After the whole process, the human activity recognition model comes into play where the measurements of the gyro-sensor and accelerometer are calculated for better results. The notations of the equations are given in the Table 1.

Table 1. Notation Table

Notation	***Meaning***
L	*Cost function*
$H(\cdot,\cdot)$	*Cross Entropy*
$y, F(X)$	*Distributions of Cross Entropy*
λ_j	*Reputation of the regularization function*
P_j	*Regularization function*

The HAR is classified into $K = 2$ because of the two sensors i.e., accelerometer and gyro-sensor. To reduce the cost of the whole process a cost function is used which has been taken from the [16]. The cost function has been given as follows:

$$L = H\big(y, F(X)\big) \tag{1}$$

where $H(\cdot,\cdot)$ defines cross-entropy among two distributions. However, the model trained with above equation exhibit higher false positive; thus, impacting overall accuracies of the model. Hence, due to this problem, this paper introduces a new cost function for the training objective as given in below equation

$$L = l(F(X), y) + \sum_j \lambda_j \, P_j \tag{2}$$

Here the $l(\cdot,\cdot)$ stands for the loss function, P_j is the regularization or consequence function, and λ_j hold the reputation of the regularization or the consequence function. The simulation analysis provides evidence that the HAR technique is more accurate

4. RESULTS AND DISCUSSIONS

In this section, the UCI HAR dataset was used for the evaluation of our model. In this dataset 30 individual performed six normal day-day-life activities and the instructor evaluated these results. There were 10,299 different activities which was divided for the testing and training our model. In Figure 3. the data representation of the UCI-HAR dataset has been shown. The different individuals performed different activities and had different outcomes for that particular activity.

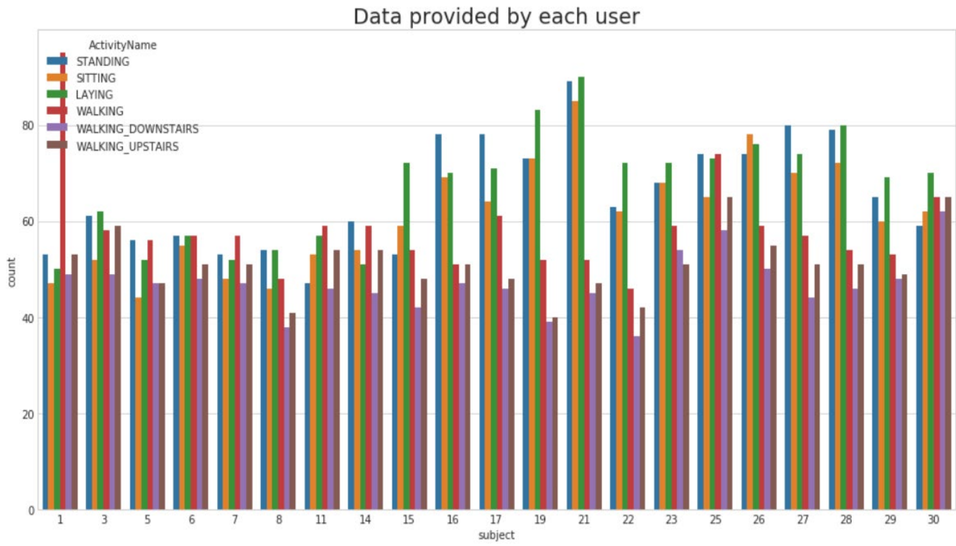

Figure 4. Data representation of different individuals

In Figure 5. the number of data points of each activity has been represented using a graph. The activities performed by the different individuals have been grouped and calculated for each activity.

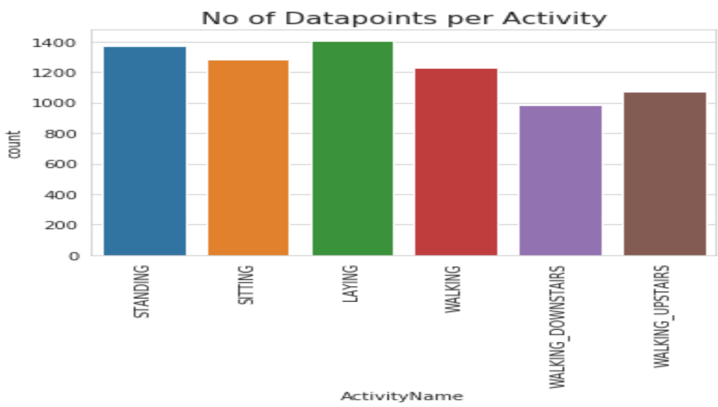

Figure 5. Data representation of every activity of different individuals

In Figure 6. the connection between the moving and the stationary activity has been represented using the graph. From this figure it can be seen that the frequency is high for the stationary activity and for the motion activity the frequency is relatively less.

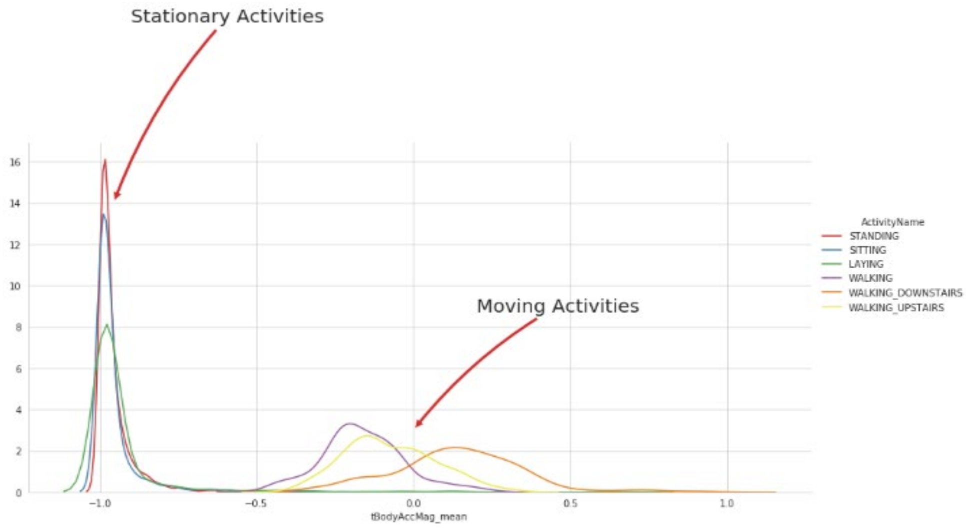

Figure 6. Data representation of moving and stationary activities

In Figure 7 the enhanced representation of data for the moving and stationary activities has been shown.

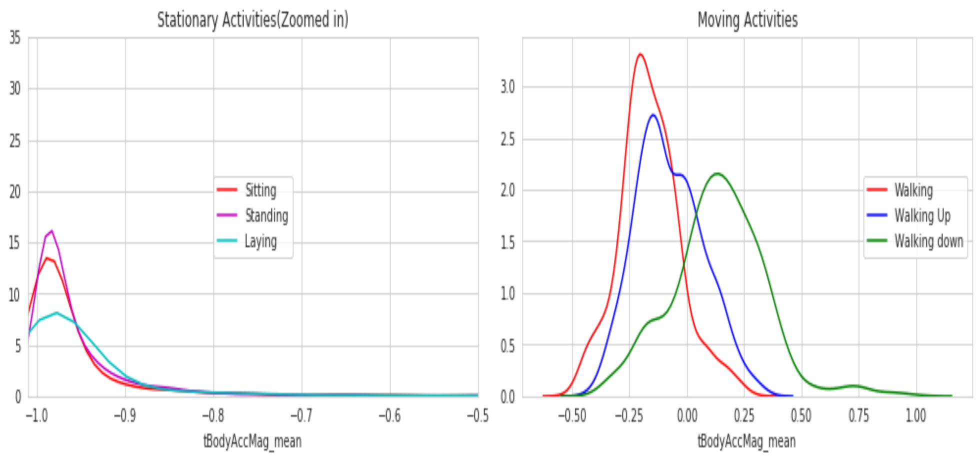

Figure 7. Data representation of moving and stationary activities

In Figure 8 the representation of the data for the acceleration magnitude of various activities has been plotted.

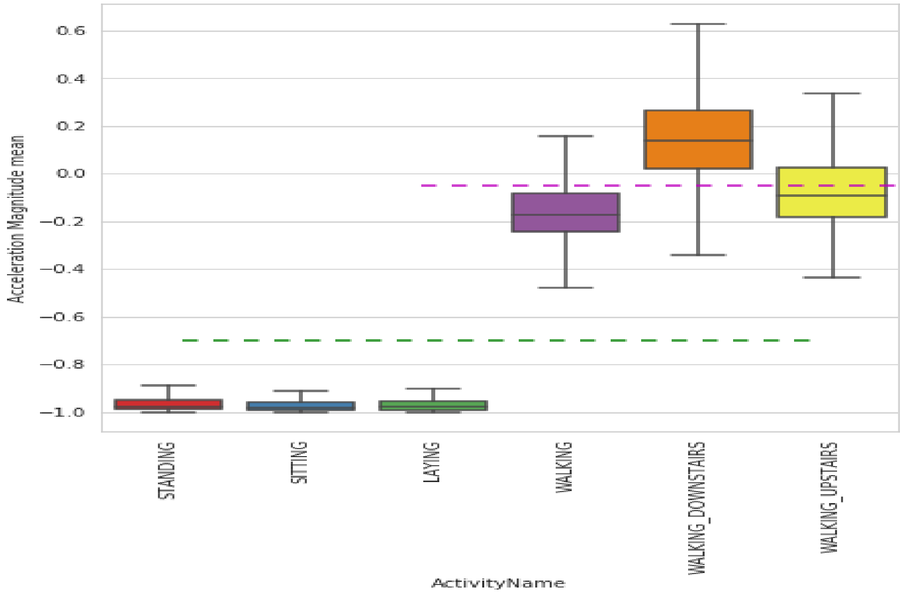

Figure 8. Data representation for the acceleration magnitude of various activities

In Figure 9 and 10, the representation of data for the various activities on the basis of Angle X and Angle Y have been represented.

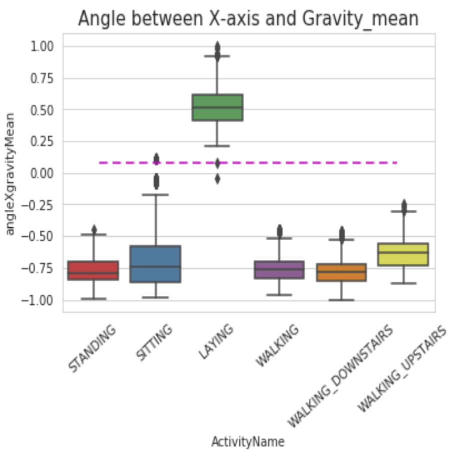

Figure 9. Data representation of various activities for Angle X

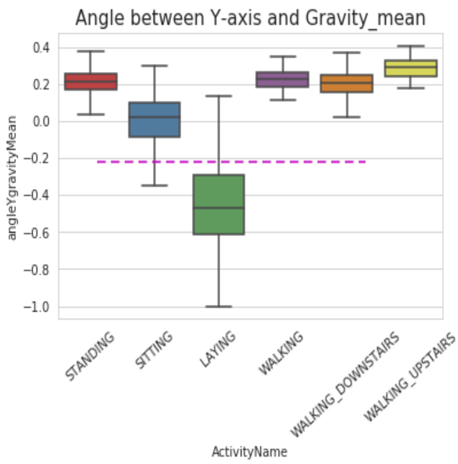

Figure 10. Data representation of various activities for Angle Y

The validation and training loss outcome attained by our model is presented in Figure 11 and the graph for validation and training accuracies outcome for our model is presented in Figure 12.

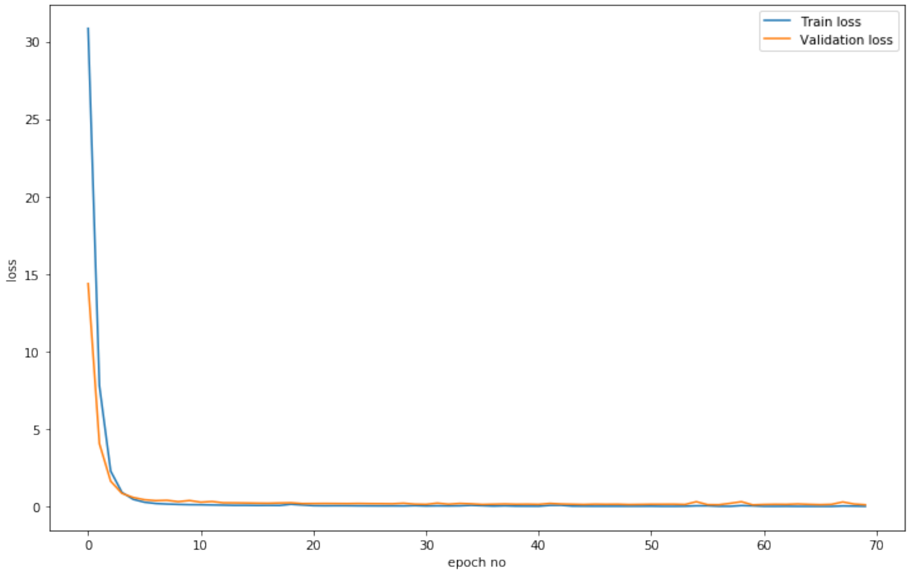

Figure 11. Data representation for the validation and training loss outcome in HAR

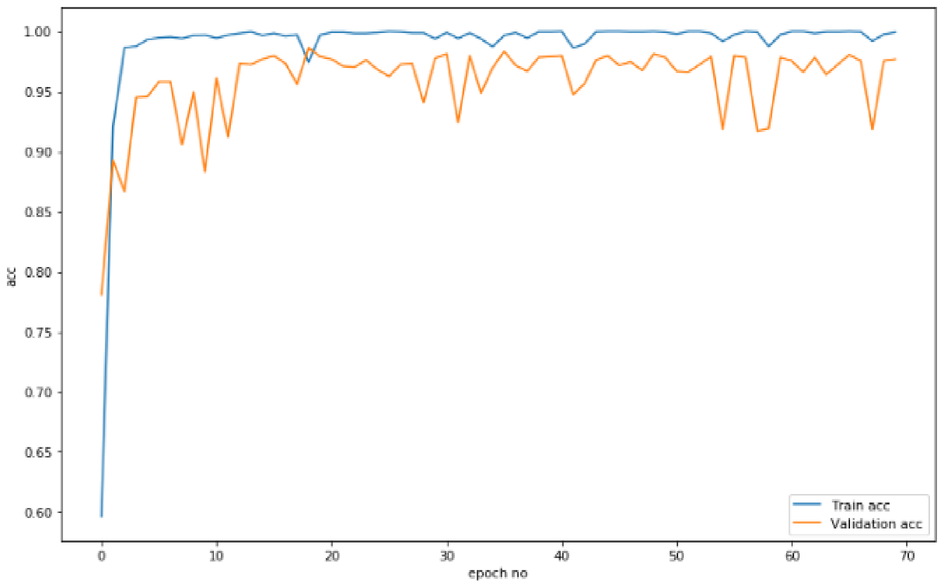

Figure 12. Data representation for the validation and training accuracies outcome in HAR

In Figure 13 the confusion matrix for the HAR for our model is shown. From this confusion matrix we have obtained the recall, precision, accuracy and F1-score.

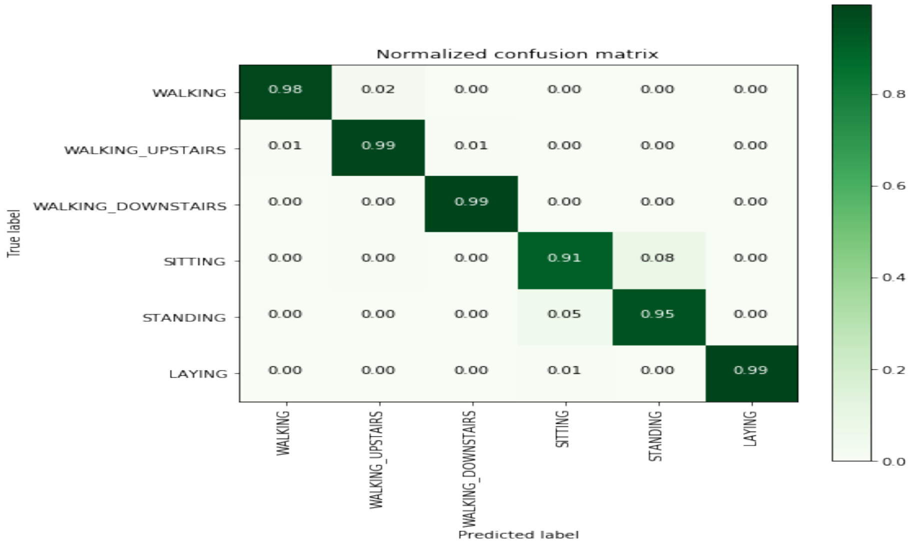

Figure 13. ROC Confusion Matrix for the CNN-GRU model

The Table 2 shows the attained results from the confusion matrix. It can be seen that our model CNN-GRU based HAR gives better performance when compared with the existing system.

Table 2. Comparison of the existing system and proposed system

	Accuracy	*Precision*	*Recall*	*F1-Score*
HARS, 2020 [16]	*95.99*	*96.04*	*95.98*	*96.01*
Our model	***96.79***	***96.93***	***96.78***	***97.82***

5. CONCLUSION

In this paper, we have proposed our CNN-GRU model for the prediction of HAR which uses the deep learning concepts to predict the different activities and the sudden change in the activities performed by the individual. This method contains an algorithm to recognize the different activities performed by the different individuals. The sensors recorded a wide range of motions and speeds along three axes, providing the dataset with a wealth of information. In order to forecast a person's typical behaviors, we first used our CNN-GRU model to standardize the data and then extracted all the relevant characteristics and properties. The accuracy of our model showed better performance when compared with the existing model. This model can be used for the healthcare purpose to identify the patient's activities who have less mobility and can be used on the patients who have to go dialysis. Moreover, the advantage of the system are as follows: this system can monitor health, discover activity patterns, detect activity and improve the wellbeing of any individual. The disadvantage of this system is that this system if the training values are not correct, then this system fails to predict the activity which is being performed.

For the future work, this model can be tested on different datasets to provide more accurate and precise results and also reduce the computation time. Our model can also be used to solve the regression problems.

REFERENCES

[1] Corpuz, R. S. A. An application method of long short-term memory neural network in classifying English and Tagalog-based customer complaints, feedbacks, and commendations. *International Journal of Information Technology and Security*, vol. 13, no.1, 2021, pp. 89-100.

[2] Akhmedova, Sh., Stanovov, V., Kamiya, Yu. Solving the human status detection problems by automatically generated neuro-fuzzy classifiers. *International Journal on Information Technologies and Security*, vol. 13, no.2, 2021, pp. 31-42.

[3] Irshad, S., Ahsan, K., Khan, M.A., Iqbal, S., Hussain, M.A., Shafiq, F., Emad, S.M.. Context-aware wearable system for animals – an exploration and classification. *International Journal on Information Technologies and Security*, vol.13, no.3, 2021, pp. 3-14.

[4] Reyes-Zárate, G.G., Cavazos, I.G. Ergonomic assessment using KINECT. *International Journal on Information Technologies and Security*, vol. 12, no.4, 2020, pp. 13-22.

[5] Khalifa, S., Lan, G., Hassan, M., Seneviratne, A., Das, S.K. HARKE: Human activity recognition from kinetic energy harvesting data in wearable devices. *IEEE Transactions Mobile Computing,* vol. 17, no. 6, Jun. 2018, pp. 1353-368.

[6] Tufek, N., Ozkaya, O. A comparative research on human activity recognition using deep learning. *Signal Processing Communication Application Conference*, Apr. 2019, pp. 1-4.

[7] Jain, A., Kanhangad, V. Human activity classification in smartphones using accelerometer and gyroscope sensors. *IEEE Sensors Journal*, vol. 18, no. 3, Feb. 2018, pp. 1169-1177.

[8] Gaikwad, N.B., Tiwari, N. B. ., Keskar, V, Shivaprakash, N.C. Efficient FPGA implementation of multilayer perceptron for real-time human activity classification. *IEEE Access*, vol. 7, 2019, pp. 26696-26706.

[9] Tufek, N., Yalcin,M., Altintas, M., Kalaoglu, F., Li, Y., Bahadir, S.K. Human action recognition using deep learning methods on limited sensory data. *IEEE Sensors Journal*, vol. 20, no. 6, Mar. 2020, pp. 3101-3112.

[10] Sukor, A.S.A., Zakaria, A., Rahim, N.A. Activity recognition using accelerometer sensor and machine learning classifiers. *IEEE 14th International Colloquium Signal Process Application (CSPA)*, Mar. 2018, pp. 233-238.

[11] Aljarrah, A.A., Ali, A.H. Human activity recognition using PCA and BiLSTM recurrent neural networks. *International Conference Engineering Technology its Application (IICETA)*, 2019, pp. 156-160.

[12] Zhang, H., Xiao, Z., Wang, J., Li, F., Szczerbicki, E. A novel IoT-perceptive human activity recognition (HAR) approach using multihead convolutional attention. *IEEE Internet Things Journal*, vol. 7, no. 2, 2020, pp. 1072-1080.

[13] Zebin, T., Scully, P.J., Peek, N., Casson, A.J., Ozanyan, K.B. Design and implementation of a convolutional neural network on an edge computing smartphone for human activity recognition. *IEEE Access*, vol. 7, 2019, pp. 133509-133520.

[14] Yu, S., Qin, L. Human activity recognition with smartphone inertial sensors using bidir-LSTM networks. *International Conference Mechanical Control Computing Engineering (ICMCCE)*, Sep. 2018, pp. 219-224.

[15] Dhanraj, S., De, S., Dash, D. Efficient smartphone-based human activity recognition using convolutional neural network, *International Conference Information Technology (ICIT)*, Dec. 2019, pp. 307-312.

[16] Yen, C.-T., Liao, J.-X., Huang, Y.-K. Human daily activity recognition performed using wearable inertial sensors combined with deep learning algorithms. *IEEE Access*, vol. 8, 2020, pp. 174105-174114

[17] Xia, K., Huang, J., Wang, H. LSTM-CNN architecture for human activity recognition. *IEEE Access*, vol. 8, 2020, pp. 56855-56866.

Information about the authors:

Srikanta Swamy Pushpalatha is a research scholar from DR. Ambedkar Institute Technology, Bengaluru. She has completed her M. Tech from VTU. Presently pursing her research work in Machine Learning Techniques, under the guidance of Dr. Shrishail Math, Professor, SKIT, Bengaluru. She is life member for many professional bodies like IE (Institute of Engineers), ISSS etc. She can be contacted at email: pushpalat043@gmail.com.

Dr. Shrishail Math, Professor in Sri Krishna Institute of Technology, Bengaluru. Presently he is guiding many students in the research areas like Computer Vision, Computer Forensics and Information Security. He has presented and published many papers in many international journals. He is a life member for many bodies like ISTE (Indian Institute for Technical Education), IAENG (International Association of Engineers), IACSIT (International Association of Computer Science and Information Technology). He can be contacted at email: sri_math@yahoo.com.

Manuscript received on 01 August 2022
First revision – 14 September 2022
Final revision – 17 October 2022