IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol. 8, Issue.11, November 2020 E-ISSN: 2347-2693

Comparative Study of the Deep Learning Neural Networks on the basis of the Human Activity Recognition

Saurav Singla[1*], Anjali Patel[2]

[1]Independent Researcher, Gurgaon-122011, India
[2]Department of Electronics and Communication, University of Allahabad, Prayagraj-211002, India

[*]*Corresponding Author: sauravsingla08@gmail.com, Tel.: +91-9958793952*

DOI: https://doi.org/10.26438/ijcse/v8i11.2732 | Available online at: www.ijcseonline.org

Received: 11/Nov/2020, Accepted: 19/Nov/2020, Published: 30/Nov/2020

Abstract - The Human Activity Recognition using the Signal produced by the Sensors have a number of applications in field of the fitness and health. The Human activities are recorded with the help of the various types of the sensor embedded in a wearable device or in a smartphone. There are many research works have been done for the Human Activity Recognition using the machine-learning as well as deep-learning models, but there is requirement to find out that which model is more efficient for a specific dataset, for which the comparative study of the model comes in mind. In this research paper the comparative study of three most efficient Deep Learning models LSTM-RNN, GRU-RNN and CNN has been performed on the most famous dataset 'Human Activity Recognition Using Smartphones Data Set' present at UCI machine-learning repository. 'LSTM-RNN' is abbreviated for 'Long Short-Term Memory-Recurrent Neural Network' is an updated version of the recurrent neural network based on the concept of back-propagation, is capable of remembering the dependencies for comparatively longer time-span. 'GRU-RNN' is abbreviated for 'Gated Recurrent Units-Recurrent Neural Network' is also an updated version of the recurrent neural network based on the concept of back-propagation, with fewer parameters than LSTM-RNN. 'CNN' is abbreviated for 'Convolutional Neural Network' is a feed forward Neural network using Convolutional layers for feature-extraction and fully-connected layer for classification.

Keywords—Human Activity Recognition (HAR), LSTM-RNN, GRU-RNN, CNN

I. INTRODUCTION

At the present time, there are many portable mobile devices having a variety of in-built sensors as GPS, accelerometers, gyroscopes, finger-print sensor, ambient-light sensor etc, which can be used in the recognition of the human activities.
There are a number of applications of the human activity recognition in the medical field.

Initial research work on the human activity recognition has been done on the basis of the devices attached on the body, later on it shifted to the smartphone sensors.

An experiment has been performed under UCI machine learning repository, and prepared a dataset which is present under the name 'Human Activity Recognition Using Smartphones Data Set' has been used for the comparative study of the Deep Learning Models in this paper.

The experiment was performed on a group of 30 person of the age in range of 19-48 years. The six activities [Walking, Walking-Upstairs, Walking-Downstairs, Sitting, Standing, Laying] were performed by each person having a smartphone (Samsung Galaxy S II) near their waist. Using the embedded accelerometer and the gyroscope, 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50 Hz were recorded. The dataset used is human activity recognition using smartphones.

In the paper 'LSTM-RNN' stands for 'Long Short-Term Memory-Recurrent Neural Network', 'GRU-RNN' stands for 'Gated Recurrent Units-Recurrent Neural Network', 'CNN' stands for 'Convolutional Neural Network' and 'HAR' stands for 'Human Activity Recognition'.

In this paper, I have performed the comparative study of three Deep Learning Models LSTM-RNN, GRU-RNN, CNN on the Human Activity Recognition.

The paper has been divided into 6 sections. Section I is Introduction, Section II is about the Related Works, Section III describes the Terminology and Calculation used in detailing and applying the different models, Section IV is Methodology in which the three deep-learning models applied for this study are described in detail, Section V is about Result and Discussion and the last VI Section is Conclusion and Future Scope.

II. RELATED WORK

In the past years many research works have been performed for the Human Activity Recognition using different machine-learning models and deep-learning models. Some of them are also based on the comparative study of the models for HAR.

Machine-Learning (Compliment Of Deep Learning) Based HAR

A study on performance evolution of the machine learning model algorithms for Human Activity Recognition using UCI data-set has been performed, by analysing some previous literature works and applying the data- collection, feature extraction, feature reduction on the data- set and trained and tested the Gaussian Naïve Bayes, K-Nearest Neighbours Classifier, Decision tree, Random-forest and Support Vector Machine Classifier and represented a comparative graph. This study concluded that Support Vector Machine showing the best performance out of those trained models on the basis of the accuracy and f-1 score. The limitations of this study were to perform prediction using the deep- learning models like CNN, RNN, LSTM and compare it with those machine-learning algorithms [1].

A survey has been done for the comprehensive overview of the performance of the trending machine learning and data-mining techniques by studying and analysing around 25-30 literature works based on the different datasets of the home activity recognition, human activity recognition and sport activity recognition. Then, performed a comparative study of the transfer learning, active learning, deep learning and semantic learning on the basis of previous research work and studies. This survey concluded that every type of learnings had its own advantage and dis-advantage but the deep-learning based models are in trend and outstanding-performance [2].

Deep-Learning Based HAR

A comparative study on the deep learning models for the real time human activity recognition has been performed by training and testing the CNN, LSTM, Bidirectional-LSTM and Support Vector Machine(SVM) on the UCI-dataset and PAMAP-2 dataset and concluded that the CNN model is best in the case of both datasets by obtaining the best accuracy, 93% in the case of UCI-dataset and 91% in the case of the PAMAP-2 dataset and the second best model in the case of the UCI-dataset is SVM and in the case of PAMAP-2 is the LSTM. The limitations and future scope of this study are more optimize the structure of the neural networks for the experiments and perform the comparative and experimental study over more other neural networks [3]

A computer-vision based study of the Human Activity Recognition has been performed on the two video datasets UCI-50 dataset and HMDB-51 dataset having the labels Clap, Climb, Walking, Cartwheel, Eat, Catch, Push-ups and Wave by training two models Convolutional Neural Network(CNN) and Hidden Markov Model(HMM) on 70% of the datasets and testing on the 30% of the dataset and compared the performance of the both on the different labels and concluded CNN performed most efficiently than the HMM models which was optimized further in that study than the previous implemented models[4].

Other Miscellaneous Work

A review on the image-processing based Human Activity Recognition has been performed, by doing the survey on scientific literature having research in HAR using various techniques of machine learning and deep learning based on image-processing dataset. This review represents the comparison of the performance of different approaches on the different datasets of image based Human Activity Recognition done in the various already available research work. This review achieved its aim of stressing out the need of automating the processing and classification with the purpose of recognizing human activities from an image dataset [5].

A detailed study has been performed on the sensor based HAR discussing the generalized architecture composing of Data Acquisition and Activity Recognition. For this study an analysing of the Online System and Offline system has been performed and a comparative analysis of the performance of the Decision Tree Classifier, Tree Classifier, Random Forest Classifier, Artificial Neural Network (ANN) and the RNN models for various three type of the feature selections [6].

Apart from all these works there are many other research works have been done related to this study.

III. TERMINOLOGY AND CALCULATION

The basic terminology, used in the paper are here in detail:

- **Activation Function:** Let j be a neuron. The activation function is defined as:

$$a_j(t) = \text{fact}\ (\text{net}\ j(t),\ a_j\,(t - 1),\ \Theta_j).$$

- **RELU activation function:** This activation function can be defined as:

$$\text{Max}\ (0, z)$$

- **Tanh activation function:** Tanh squashes a real-valued number to the range [-1, 1]. It's non-linear. But unlike Sigmoid, its output is zero-centered.

$$\text{Tanh}(z) = 1/\ (1 + e^{-z/T})$$

- **Sigmoid Function:** It can be defined as:

$$S(z)=1/(1+e^{-z})$$

- **Soft-max Function:** It can be defined as:

$$\text{Soft-max}\ (z) = e^{Z}{}_{i}/\textstyle\sum^{k}_{j=1} e^{z}{}_{j}$$

- **Accuracy:** It is the most relevant measure for the performance of any machine learning model. It is the ratio of the correctly predicted observation to the total observation.

Accuracy = (TP+TN)/(TP+TN+FP+FN)

- **Precision:** It is the ratio of the correctly predicted positive observations to the total predicted positive observations.

Precision = TP/(TP+FP)

- **Recall:** It is the ratio of the correctly predicted positive observation to the total observation in the actual class.

Recall = TP/(TP+FN)

- **F1-Score:** It is the weighted average of the precision and recall.

F1-score = 2*(Recall*Precision)/ (Recall + Precision)

- Where:

  TP = True Positive
  TN = True Negative
  FP = False Positive
  FN = False Negative

IV. METHODOLOGY

LSTM-RNN

LSTM-RNN method is one of the most efficient method to deal with the *time-series prediction problems*. This method was proposed by in 1997 [Sepp and Jurgen] to deal with the vanishing gradient problem.

The LSTM-RNN is a special version of the Recurrent Neural Network (RNN). Lets first focus on the RNN. RNN is neural network is based on the back-propagation algorithm applied at every time stamp. The RNN method uses the prediction at time t-1 and the new information at time t to do the prediction at the time t. This concept can be easily understood with the following diagram.

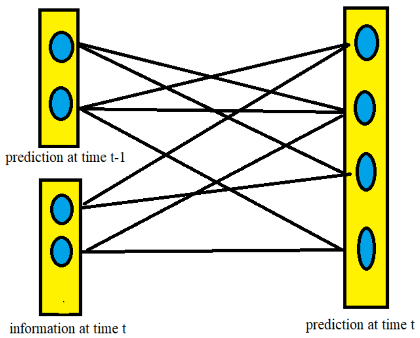

Figure 1. Prediction Method of RNN

Figure 1 is representing that how the Recurrent Neural Network (RNN) Works which has been discussed above.

But in this method of RNN there were the limitation of learning from the dependencies of the few back time prediction and problem of the Vanishing Gradient Problem.

To over-come these limitations the LSTM-RNN was introduced which is capable for learning from the long-term dependencies.

The Architecture of the LSTM-RNN comprises of some special memory-blocks along with the recurrent hidden layers. Each memory block contains self-connected memory cells which store the contemporary state of the network along with special type of the multiplicative units called as gate in order to control the flow of the information. The latest updated architecture of the of the LSTM-RNN contains three gates, but the initial architecture contained only two gates, an input gate for controlling the flow of input activations into the memory cell and an output gate for controlling the output flow of cell activations into the rest of the network. Later on, a forget gate was added which scales the internal state of the cell before adding it as input to the cell through the self-recurrent connection of the cell, therefore adaptively forgetting or resetting the cell's memory.

The advantages of LSTM-RNN is that by increasing the input threshold, forgetting threshold and output threshold, the weight of the self-loop is changed such that the integral scale at different times can be dynamically changed when the model parameters are fixed. Thus, it solves the problem of vanishing gradient or gradient expansion.

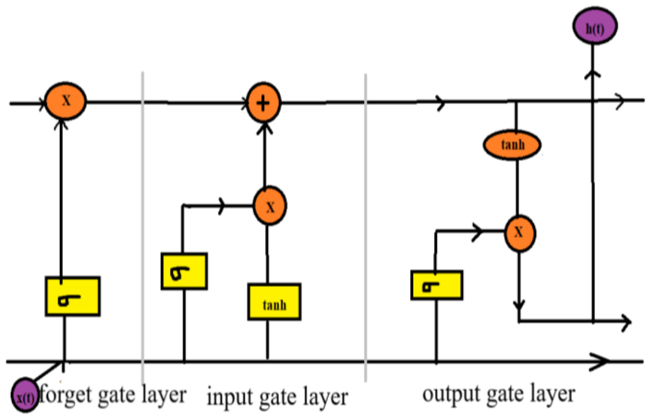

Figure 2. Architecture of LSTM-RNN unit

Figure 2 is representing the architecture of the LSTM-RNN model unit which is comprising of three gate layers.

GRU-RNN

The Gated Recurrent Units (GRU-RNN) is a variant of the Recurrent Neural Network, introduced in 2014 on the concept of the encoder and the decoder. The architecture of it is based on the model of the LSTM-RNN, consists of hidden states containing two gates, the update gate which controls that how much information from the previous hidden state will carry over the current hidden state and the reset gate which effectively allows the hidden state to drop any information that is found to be irrelevant in future. The mechanism of the gate is that the weights corresponding to these gates are updated using the backpropagation through time (BTT) stochastic gradient descent in order to minimize a loss/cost function.

Further three gate-variants of the GRU-RNN has been introduced as GRU-1 in which each gate has been updated only using the previous hidden state and the bias, GRU-2 in which each gate has been computed only using the previous hidden state and GRU-3 in which each gate has been computed only using the bias [6]. In the experiment regarding this paper the original GRU-RNN model has been used.

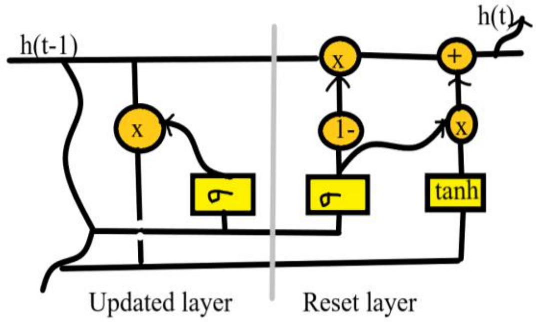

Figure 3. Architecture of GRU-RNN unit

Figure 3 is representing the architecture of the GRU-RNN model unit which is comprising of three gate layers.

CNN

The CNN contains at least one temporal convolution layer, one pooling layer and at least one fully connected layer prior to the top-level soft-max group. The temporal convolution layer corresponds to a convolution of the input sequence with different kernels (feature maps). Subsequent max-pooling is looking for the maximum within a region of width mw and corresponds to a subsampling, introducing translational invariance to the system. Here, two convolutional layers have been used as shown in the figure.4. The output of each max-pooling layer is transformed using a RELU activation function.

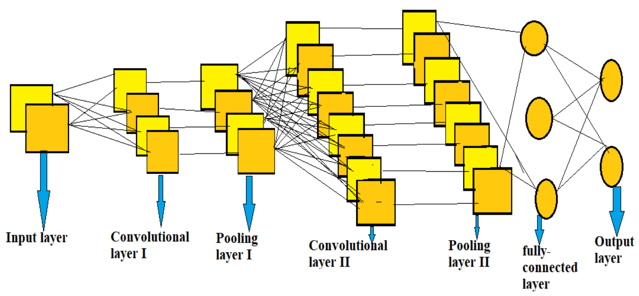

Figure 4. Architecture of CNN

Figure 4 is representing that architecture of the CNN model which is comprising of two convolutional layers along with input and output layer.

V. RESULTS AND DISCUSSION

The above discussed methodology has been train on the 70% of data-set discussed in the 'Section I' by using KERAS and then tested on the remaining 30% of the dataset found the following results:

LSTM-RNN

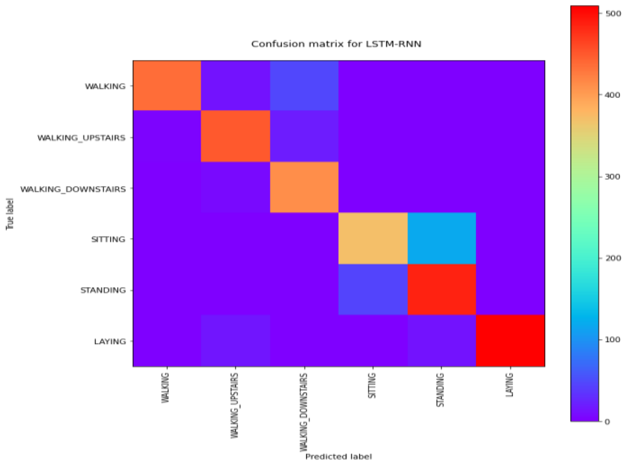

Figure 5. Confusion Matrix for LSTM-RNN

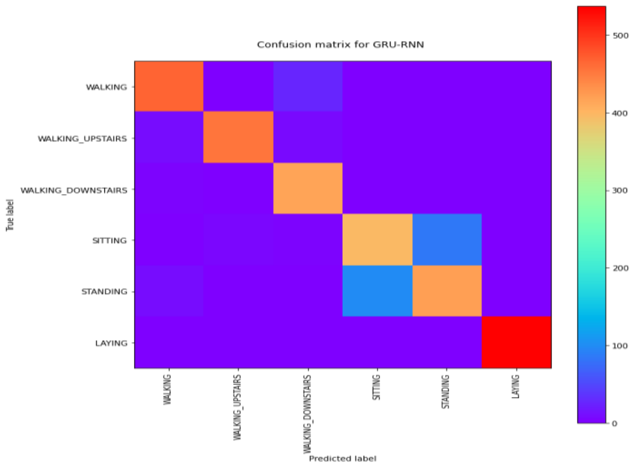

Figure 6. Confusion Matrix for GRU-RNN

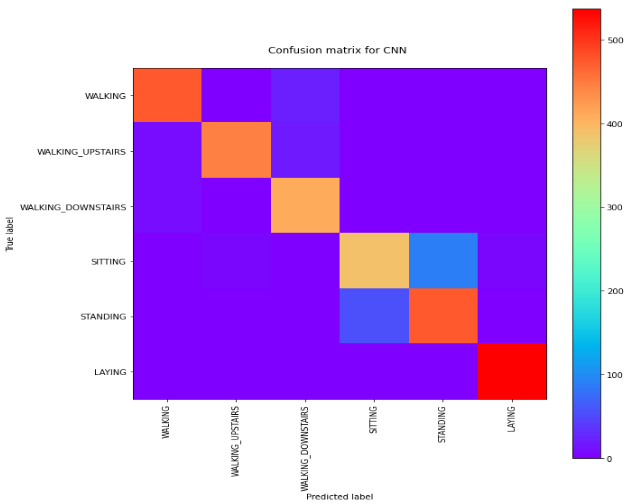

Figure 7. Confusion Matrix for CNN

Figure 5, 6, 7 are representing the confusion matrices formed with true label at x-axis and predicted label at y-axis obtained by the LSTM-RNN, GRU-RNN, CNN models respectively.

Table 1. Classification Report for LSTM-RNN

No	precision	recall	f1-score	support
0	0.99	0.88	0.93	496
1	0.93	0.93	0.94	471
2	0.86	0.86	0.92	420
3	0.89	0.89	0.82	491
4	0.79	0.79	0.85	532
5	1.00	1.00	0.97	537

The Table 1 is representing the classification report for the LSTM-RNN model comprising of the precision, recall, f1-score, support parameters which are discussed in the Section III.

Table 2. Classification Report of the GRU-RNN

No	precision	recall	f1-score	support
0	0.94	0.97	0.95	496
1	0.91	0.94	0.92	471
2	0.98	0.98	0.98	420
3	0.82	0.82	0.82	491
4	0.84	0.83	0.84	532
5	1.00	0.95	0.97	537

The Table 2 is representing the classification report for the GRU-RNN model comprising of the precision, recall, f1-score, support parameters which are discussed in the Section III.

Table 3. Classification Report for CNN

No	precision	Recall	f1-score	support
0	0.98	0.94	0.96	496
1	0.98	0.93	0.95	471
2	0.88	1.00	0.94	420
3	0.77	0.78	0.78	491
4	0.84	0.78	0.81	532
5	0.96	1.00	0.98	537

The Table 3 is representing the classification report for the CNN model comprising of the precision, recall, f1-score, score parameters which are discussed in the Section III.

Table 4. Aggregate Accuracy of the Models.

No.	Model	Accuracy
1	LSTM-RNN	90%
2	GRU-RNN	91%
3	CNN	93%

Table 4 is representing the Aggregate Accuracy obtained by the different models trained and test in this study.

The Experiments performed for proposed model are showing the above results from which it can be concluded that the CNN model giving most efficient performance.

Results obtained from the various previous works:

Table 5. Aggregate Accuracy of the Models discussed in related work

	Models	Accuracy
1	CNN	92%
2	LSTM & BLSTM	89%
3	MLP	86%
4	ANN	94%

Table 5 is showing the results summary of the deep learning models discussed in the related work (section II).

The figures and table mention above are showing that the models proposed in this paper are performing more efficiently in comparison to the some previously done works and studies. The LSTM-RNN model trained and tested for this paper is giving the accuracy of 90% which is more efficient than the LSTM and BLSTM models discussed in the related work giving accuracy of 89% and MLP which is giving the accuracy of the 86%. The second proposed model GRU- RNN has not been discussed in the related work, gives a better accuracy of 91%. These two proposed models are less efficient than some previously trained model as CNN (92% accuracy) and ANN (94% accuracy). The third proposed model CNN is more efficient than CNN model of the previous studies but lass efficient than the ANN.

VI.CONCLUSION AND FUTURE SCOPE

According to the above study for the taken dataset the all the three models have given the highest accuracy in range 94-95 but the aggregate accuracy obtained by the LSTM-RNN is 90, GRU-RNN is 91 and by the CNN is 93. As the Accuracy the most relevant parameter of the performance and the highest accuracy of all three models is almost same but the aggregate accuracy of the CNN is Highest (93%).

From the above studies and comparison, we can conclude that the model proposed in this paper are more efficient than many of the previously proposed models but not the most efficient as the ANN model discussed in the previously performed related work.

The future scope is to maximize the accuracy by applying the different deep learning and machine-learning models and approaches like optimization and hyper-tunning of the model.

REFERENCES

[1] S. S. Anju, K. V. Kavitha, "*Performance Evolution of Varoius Machine Learning Technique for Human Activity Recognition using Smartphone*", Vol.**7**, Issue.**8**, pp.**316-319, 2019.**

[2] S. R. Ramammurthy, N. Roy, *"Recent trends in machine learning for human activity recognition- A Survey"*, Wiley Interdisciplinary Reviews: Data Minining and Knowledge Discovery, **2018**, doi: 10. 1002/widm. 1254.

[3] S. Wan, Q. Lianyong, X. Xiolong, T. Chao, G. Zonhua, *"Deep Learning Models for Real-time Human Recognition with*

Smartphones", Mobile Networks and applications, **2019**, doi: 10. 1007/S11036. 019- 01445- x.

[4] G.A. Kani, P. Geetha, A. Gomathi, *"Human Activity recognition using deep learning with Ggradient Fused Handcrafted Features and Categorization based on Machine Learning Technique",* International Journal of Computer Sciences and Engineering, Vol.**6**, Issue.**7,** pp.**1-7, 2018.**

[5] N. Geeta, E.S. Samundeeswari, "*A Review on Human Activity Recognition System"*, International Journal of Computer Sciences and Engineering, Vol.**6**, Issue.**12,** pp.**825-829, 2018.**

[6] M. Badshah, *"Sensor-based Human Activity Recognition using Smartphones",* Sensors, **2019,** doi: 10.31979/etd.8fjc-drpn.

[7] R. Dey, F.M. Salem, "*Gate-Variants of Gated Recurrent Unit (GRU) Neural Networks"*, journal of IEEE Acess,**2017,** doi: 10.1109/MWSAS.2017.8053243.

[8] W. Jiang, Z. Yin, and., " *Human activity recognition using wearable sensors by deep convolutional neural networks.",* In Proceedings of the 23rd ACM international conference on Multimedia (pp**. 1307-1310), 2015.**

[9] H. Cho, S.M. Yoon, "*Divide and Conquer based 1D CNN Human Activity Recognition using test data Sharpening",* Sensors, **2018, 1055,** doi:10.3390/s 18041005.

AUTHORS PROFILE

Saurav Singla is a Senior Data Scientist and a Machine Learning Expert. He has fifteen years of comprehensive experience in statistical modeling, machine learning, natural language processing, deep learning, and data analytics. He has a Master of Science from the University of Westminster.

He has been recognized for maximizing performance by implementing appropriate project management tools through analysis of details to ensure quality control and understanding of emerging technology.

Outside work, Saurav volunteers his spare time for helping, coaching, and mentoring young people in taking up careers in the data science domain. He has created two courses on data science, with over twenty thousand students enrolled in it. He regularly authors articles on data science.

Anjali Patel is pursuing Bachelors of technology in Computer Science and Engineering from University of Allahabad, Prayagraj. She has done many self-guided projects in the field of Data-Science and Machine-Learning. She is a part-time junior data-scientist at Avishkaar Tech Solutions. She has been Completed Machine-Learning And Data-Analytics summer training internship at Indian Institute of Technology-Banaras Hindu University(IIT-BHU) Apart from the technical skills and education, She is an educational youtuber and has been Internshala Student Partner (ISP).

