Research Article

Human activity recognition and classification using of convolutional neural networks and recurrent neural networks

Mohammed Albaba [a] ⓘD, *Alhakam Qassab* [b] ⓘD, *Arif Yılmaz* [c,*] ⓘD

[a]*Electrical and Electronics Engineering Department, Ankara Yildirim Beyazit University, Ayvali, Kecioren, Ankara 06010, Turkey*
[b]*Computer Engineering Department, Ankara Yildirim Beyazit University, Ayvali, Kecioren, Ankara 06010, Turkey*
[c]*ODTU Campus, TUBITAK UZAY, Cankaya, Ankara 06100, Turkey*

ARTICLE INFO

ABSTRACT

Recently, by using the deep learning models, it has become easier to recognize the human activity with more accuracy than before by categorizing the activities that people are doing daily. Nowadays, with the extensive use of modern smartphones that have sensors, it has become easier to capture the data in raw format that has the movement details in three dimensions (X-Y-Z). In this paper, we utilized the open source WIreless Sensor Data Mining (WISDM) dataset which has six activities that are walking, jogging, standing, sitting, upstairs and downstairs. Each type of those activities consists of values in terms of (X, Y and Z) axes. We employed two types of deep learning algorithms that are Convolutional Neural Network (CNN) and Recurrent Neural Network - Long Short-Term Memory (RNN-LSTM). Our objective is to make a comparison between accuracy and loss after implementing the two models. We discovered that, when using the Convolutional Neural Network (CNN), the accuracy was 81%. However, the accuracy was 91% when using Recurrent Neural Network - Long Short-Term Memory (RNN-LSTM) and applying it on the same database. As a result, the Recurrent Neural Network - Long Short-Term Memory (RNN-LSTM) model outperformed the Convolutional Neural Network (CNN) model.

1. Introduction

Thanks to recent advances in sensing technologies, recognition of human activity has taken high attention. There are many applications that depend on the recognition of human activity in fields such as sports, fitness tracking, healthcare and much more [1-4]. The type of human activity could be recognized in various ways such as using cameras or sensors. In the first one, detection is made according to the video or image which is captured from the camera, second one is from the data taken from sensors built in smartphones [5]. There are different sensor types like accelerometer, sound, luminance and other sensors [6-8]. The extraction of an activity's features is rather essential as it grabs the details and information and classifies them into the right type of activities after applying some machine learning or deep learning algorithms. The more prominent features extracted from the dataset, the higher accuracy when classifying activities [9].

The extensive use of sensors like accelerometers made it easy to identify and predict the activity which the person is doing. Not long ago, real-time processing became feasible, which enabled the computation of algorithms much faster without the need of transferring data to a more powerful server [10].

Human activity classification possesses a lot of potential in healthcare applications such as assisting doctors to make the right decision through keeping track of the patients and monitoring their behaviors in daily life in order to deliver a more precise diagnosis [11].

In this paper, we utilized the open source WIreless Sensor Data Mining (WISDM) dataset [12]. The dataset contains six activities with their values for the three X, Y

and Z axes. The activities are divided into six categories that are standing, sitting, walking, jogging, ascending and descending stairs as shown in Figure 1. The three axes are as shown in Figure 2 [13].

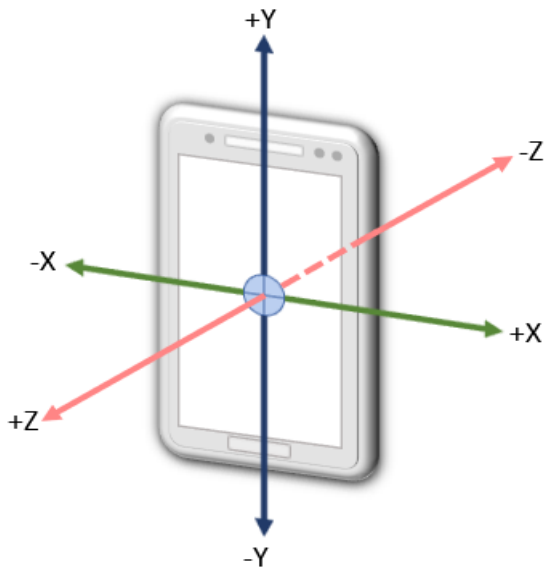**Figure 1.** Human Activity Recognition Dataset Types

Figure 2. The Three Axes Expressing Accelerometer Measurements In Smartphone (Adapted from [13])

Some neural networks in deep learning models have the ability of classifying and extracting features such as Convolutional Neural Networks (CNNs) which is a type of Deep Neural Network (DNN) [14], also Recurrent Neural Network (RNN) that has Long Short-Term Memory (LSTM) cell [8]. In this paper, human activity recognition was analyzed by performing two deep learning models that are CNN and RNN-LSTM. The aim of this paper is to find and use the best model in terms of accuracy by enhancing the recognition of daily activities to improve the healthcare case of people and monitor their actions.

2. Literature Review

2.1. Conventional Neural Network (CNN)

There are some studies that used CNN to recognize human activity. In [15], the study was done by gathering the values from the sensors in smartphones, and the

accuracy was 93.27%. In [16], depending on the raw data of the sensor, the accuracy was 91.9%.

2.2. Recurrent Neural Network (RNN-LSTM)

The study in [17] focused on RNN-LSTM model which had an accuracy of 97%. The data of human activity was gathered from a gyroscope and accelerometer in a smartphone. Another study [18] depended on camera data to classify patient human activity, the result was 92.49%.

3. Method

Our methods were trained on WISDM dataset which is a raw file that contains six different human activities as mentioned before, the values were calculated after 36 persons performed those activities. The activities have three values for each according to the axes which are (X, Y and Z). The epoch number was settled up to 15 and Python with version of 3.7 was used in both; CNN and RNN-LSTM models.

3.1. Conventional Neural Network (CNN)

CNN model in our study consists of two convolutional modules, as shown in Figure 3 which illustrates the architecture of CNN.

The data was divided into 80% for training and 20% for testing, the epoch number is 15. During the training process, the highest accuracy value was at epoch 15 and it started to lower after that. The epoch of 15 will be adopted for the final result.
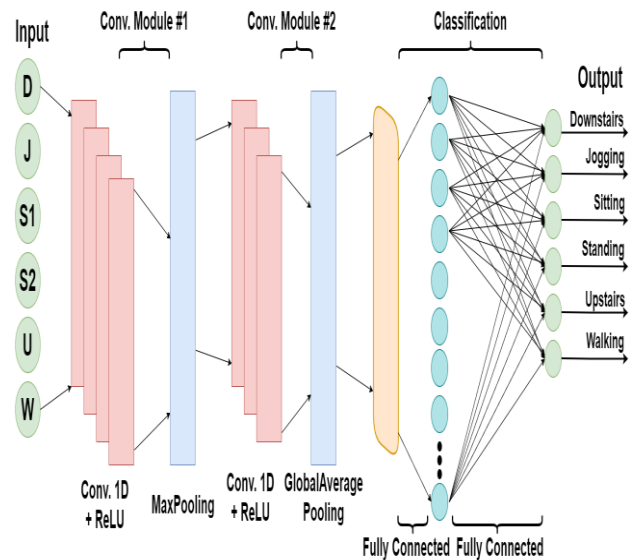

Figure 3. CNN Architecture Diagram

3.2. Recurrent Neural Network (RNN-LSTM)

Our (RNN-LSTM) model which is illustrated in Figure 4, has two layers. The model has a fully-connected layers scheme. The epoch number was settled as 15.
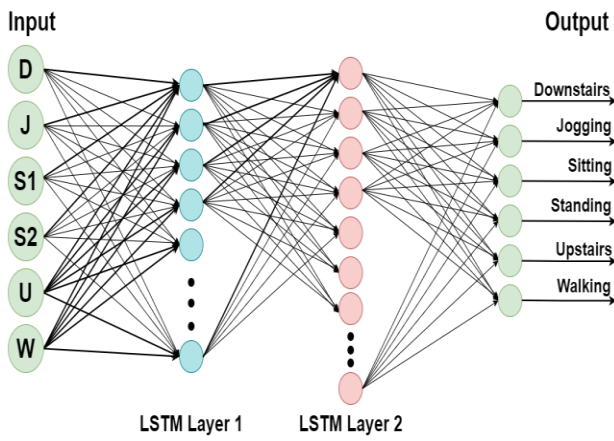
Figure 4. LSTM Architecture Diagram

4. Implementation

4.1. Conventional Neural Network (CNN)

Figure 5 shows the six activities' categories. Some libraries were used in Python such as TensorFlow and plot to draw the graphs.
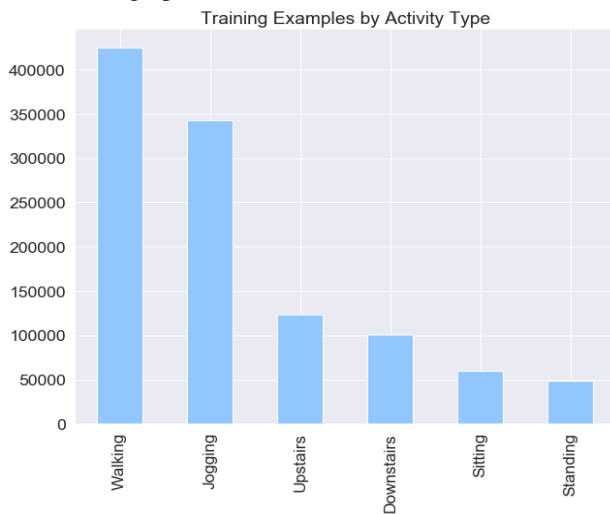

Figure 5. HAR Classification According to Activity Type

Three-axes graphs were shown for each activity, one of them is jogging activity as it is illustrated in Figure 6.
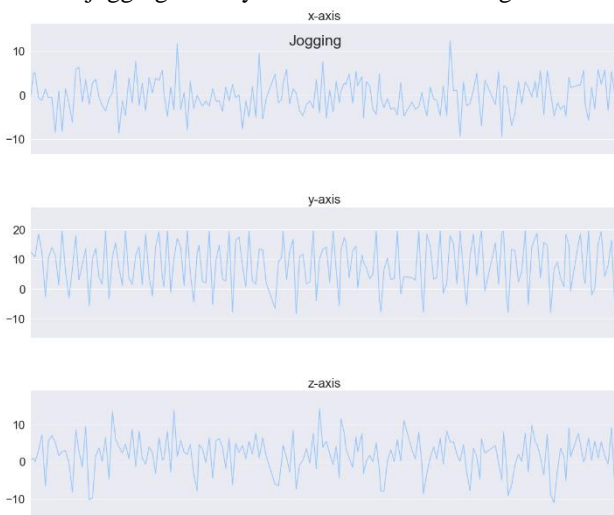

Figure 6. Jogging Activity in terms of three axes

According to Figures (7-8), it may be seen that the accuracy is 84.31% and the loss is 51.50% after applying CNN algorithm on the dataset and training the dataset for 15 epochs.

Figure 7. Learning Curve of Model Training Accuracy

Figure 8. Learning Curve of Model Training Loss

4.2. Recurrent Neural Network (RNN-LSTM)

By using it, the accuracy was 91.77% and loss was 40.89% as shown in Figures (9-10) for the 15 epochs.

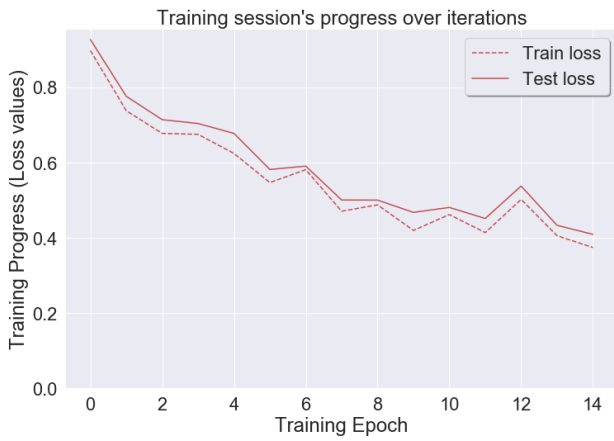**Figure 9.** Learning Curve of Model Training Accuracy

Figure 10. Learning Curve of Model Training Loss

5. Results

5.1. Conventional Neural Network (CNN)

By implementing CNN algorithm on the dataset, during the training process for every epoch, we noticed that the highest value of accuracy was at epoch 15. Therefore, we adopted this number of epochs. The whole accuracy of the testing dataset was 81%. In terms of the confusion matrix which is shown in Figure 11, it is obvious that there are some confusions happening among some activities in the prediction and true actions.
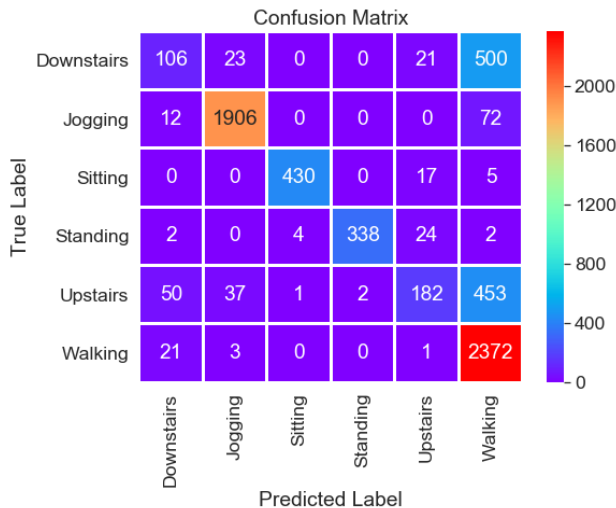

Figure 11. Confusion Matrix for Test Data

5.2. Recurrent Neural Network (RNN-LSTM)

When applying the RNN-LSTM model on the database for 15 epochs, the final test accuracy was 91%. As shown in Figure 12, the confusion matrix illustrates six activities in terms of prediction and true labels. There is a confusion between upstairs in the true label, and downstairs in the predicted label, that happened during the process, but it has a low value.
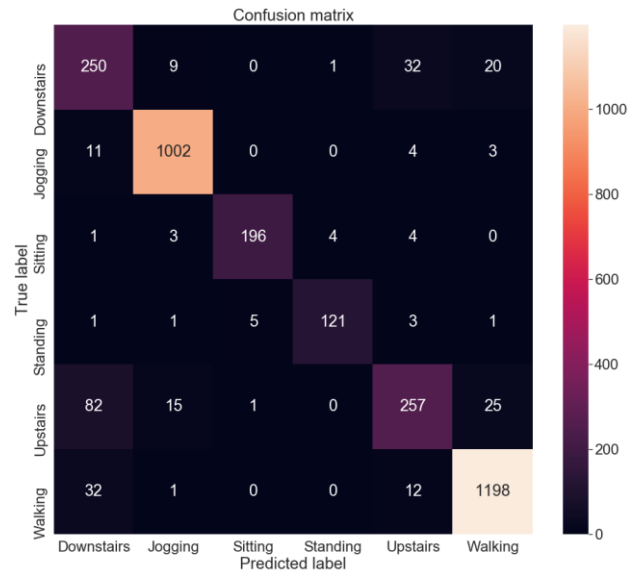

Figure 12. Confusion Matrix for Test Data s

6. Conclusion

In this work, we made a comparison between two different models in deep learning by using the WISDM dataset. We applied Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN-LSTM) models. We reached the result that the RNN-LSTM outperforms CNN as it has higher accuracy by comparing it with CNN, as their accuracy values are 91% and 81% respectively.

As a future work, a larger dataset may be used with more advanced deep learning methodologies to get more accurate and more reliable results to classify the activities with less confusion.

Author's Note

Abstract version of this paper was presented at 9th International Conference on Advanced Technologies (ICAT'20), 10-12 August 2020, Istanbul, Turkey with the title of "A Comparison of Convolutional and Recurrent Network Algorithms On Human Activity Recognition".

References

[1] C. Snehal and C. Peng-Wen and Z. Jiang and Y. Z. Joy, "Mobile Lifelogger – Recording, Indexing, and Understanding a Mobile User's Life," *Second International ICST Conference, MobiCASE 2010: Mobile Computing, Applications, and Services: Proceeding of 2nd, 2010, Santa Clara, CA, USA, October 25-28, 2010.* Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 2012. pp. 263-281. DOI: https://doi.org/10.1007/978-3-642-29336-8_15.

[2] F. Kilian and R. Daniel and T. Gerhard, "Unsupervised classifier self-calibration through repeated context occurences: is there robustness against sensor displacement to gain?," *2009 International Symposium on Wearable Computers: Proceeding of 9th, 2009, Linz, Austria, 4-7 Sept. 2009.* IEEE, 2009. pp. 77-84. DOI: 10.1109/ISWC.2009.12.

[3] W. Pang and P. Huan-Kai and Z. Jiang and Z. Ying, "SensCare: Semi-automatic Activity Summarization System for Elderly Care," *Third International Conference, MobiCASE 2011: Mobile Computing, Applications, and Services: Proceeding of 3rd, 2011, Los Angeles, CA, USA, October 24-27, 2011.* Institute for Computer Sciences, Social Informatics and

Telecommunications Engineering, 2012. pp. 1-19. DOI: http://dx.doi.org/10.1007/978-3-642-32320-1_1.

[4] W. Pang and Z. Jiang and Y. Z. Joy, "MobiSens: A Versatile Mobile Sensing Platform for Real-World Applications," *Mobile Networks and Applications,* vol. 18, pp. 60-80, 2013. DOI: https://doi.org/10.1007/s11036-012-0422-y.

[5] Z. Ming and T. N. Le and Y. Bo and J. M. Ole and Z. Jiang and W. Pang and Z. Joy, "Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors," *6th International Conference on Mobile Computing, Applications and Services (MobiCASE): Proceeding of the 6th, 6th International Conference on Mobile Computing, Applications and Services, 6-7 Nov. 2014.* IEEE, 2014. pp. 197-205. DOI: 10.4108/icst.mobicase.2014.257786.

[6] C. Diane and D. F. Kyle and C. K. Narayanan, "Transfer Learning for Activity Recognition: A Survey," *Knowledge and Information Systems,* vol. 36, pp. 537-556, 2013. DOI: https://doi.org/10.1007/s10115-013-0665-3.

[7] V. Praneeth and D. Debraj and K. D. Sajal and B. Shekhar, "A-Wristocracy: Deep Learning on Wrist-worn Sensing for Recognition of User Complex Activities," *2015 IEEE International Conference on Wearable and Implantable Body Sensor Networks (BSN): Proceeding of the 12th, Cambridge, MA, USA, 9-12 June 2015.* IEEE, 2015. pp. 1-6. DOI: 10.1109/BSN.2015.7299406.

[8] J. O. Francisco and R. Daniel, "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition," *Sensors,* vol. 16, pp. 1-25, 2016. DOI: https://doi.org/10.3390/s16010115.

[9] Z. Mi and A. S. Alexander, "A Feature Selection-Based Framework for Human Activity Recognition Using Wearable Multimodal Sensors," *6th International Conference on Body Area Networks: Proceeding of 6th, 2011, Beijing, China, November 7 - 8, 2011.* ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) Brussels, Belgium, 2011. pp. 92-98. DOI: 10.4108/icst.bodynets.2011.247018.

[10] W. P. Schalk and M. Reza, "Human Activity Recognition Using LSTM-RNN Deep Neural Network Architecture," *2019 IEEE 2nd Wireless Africa Conference (WAC): Proceeding of the 2nd, 2019, Pretoria, South Africa.* IEEE, 2019. pp. 1-5. DOI: 10.1109/AFRICA.2019.8843403.

[11] C. Wen-Hui and A. B. B. Carlos and T. Chih-Hao, "LSTM-RNNs Combined with Scene Information for Human Activity Recognition," *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom): Proceeding of 19th, 2017, Dalian, China, 12-15 Oct. 2017.* IEEE, 2017. pp. 1-6. DOI: 10.1109/HealthCom.2017.8210846.

[12] R. K. Jennifer and M. W. Gary and A. M. Samuel, "Activity Recognition using Cell Phone Accelerometers," *ACM SIGKDD Explorations Newsletter,* vol. 12, 2011. Association for Computing Machinery New York, NY, United States, 2011. DOI: https://doi.org/10.1145/1964897.1964918. WISDM (Wireless Sensor Data Mining) Project. Fordham University, Department of Computer and Information Science, http://www.cis.fordham.edu/wisdm/dataset.php [accessed date 10 May 2020]

[13] MathWorks, Accelerometer Documentation, Measure linear acceleration along X, Y, and Z axes [online] https://www.mathworks.com/help/supportpkg/android/ref/accelerometer.html [accessed date 20 July 2020]

[14] L. Yann and B. Yoshua, *The handbook of brain theory and neural networks.* MIT Press55 Hayward St. Cambridge, MA, United States, 1998.

[15] C. Heeryon and M. Y. Sang, "Divide and Conquer-Based 1D CNN Human Activity Recognition Using Test Data Sharpening," *Sensors,* vol. 18, pp. 1-24, 2018. DOI: 10.3390/s18040055.

[16] X. Li and X. Si and L. Nie and J. Li and R. Ding and D. Zhan and D. Chu, "Understanding and Improving Deep Neural Network for Activity Recognition," *11th EAI International Conference on Mobile Multimedia Communications: Proceeding of 11th, 2018, Qingdao, People's Republic of China.* EAI, 2018. pp. 1-10. DOI: http://dx.doi.org/10.4108/eai.21-6-2018.2276632.

[17] Z. U. Md and T. Jim, "Activity Recognition Using Smartphone Sensors, Robust Features, and Recurrent Neural Network," *2019 13th International Symposium on Medical Information and Communication Technology (ISMICT): Proceeding of 13th, 2019, Oslo, Norway.* IEEE, 2019. pp. 1-6. DOI: 10.1109/ISMICT.2019.8743759.

[18] S.U. Park and J.H. Park and M.A. Al-masni and M.A. Al-antari and Md.Z. Uddin and T.-S. Kim, "A Depth Camera-based Human Activity Recognition via Deep Learning Recurrent Neural Network for Health and Social Care Services," *Procedia Computer Science,* vol. 100, pp. 78-84, 2016. DOI: https://doi.org/10.1016/j.procs.2016.09.126.