International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 06 | Jun 2022 www.irjet.net p-ISSN: 2395-0072

Human Activity Recognition System

Disha Amit Nakhare [1], Palak Shah [1], Harsh Chetan Faganiya [1], Prof Uttara Bhatt[2]

[1] Undergraduate Students, Department of EXTC,Thadomal Shahani Engineering College, Mumbai, India
[2]. Assistant Professor, Department of EXTC,Thadomal Shahani Engineering College, Mumbai, India

---***---

Abstract - *Human activity recognition is a capability to interpret natural body gesture or movement via detectors and arbitrate natural exercise or action. Utmost of the human day-to-day chores can be streamlined or automated if they can be recognized via HAR system. Generally, HAR system can be additionally supervised or unsupervised. A supervised HAR system needed some previous training with constant datasets while unsupervised HAR network is being configured ground rules during development. HAR is accounted as an important element in varied scientific exploration surrounds i.e., supervision, healthcare and human computer interaction (HCI). This paper will present you an elaborate depiction and comparison of the styles, namely CNN and RNN, that can be applied to develop a Human Activity Recognition System.*

Key Words: **Human activity recognition, Convolutional Neural Network, Long Short-Term Memory, Datasets, Neural Network.**

1. INTRODUCTION

Human Activity Recognition (HAR) is one of the active exploration fields in computer vision as well as human computer relation. Still, it remains a really complicated assignment, due to unresolvable difficulties similar as detector movement, detector placement, cluttered background, and essential variability in the method activities are carried on by distinguishable human. The most generally engaged sensing technologies in HAR system regardless of the computational models or category algorithms are anatomized. The pros and cons of each sensing technology have been bandied. This paper is wrapped up with some complaints for the most sophisticated sensing technologies.

Research suggests that the physical activities led by people can affect them physically, mentally, socially and cognitively. A sedentary lifestyle can lead to a number of diseases, including but not limited to several forms of cancer, diabetes, hypertension, coronary and cerebrovascular diseases, and obesity etc. Similarly, research establishes a correlation between the amount of physical activity undertaken and mental wellbeing. With 13 percent of the total globe disease burden accounting to mental health, depression is predicted to be the leading cause for it by 2030. That can be countered with exercise as research suggests that exercise does improve the mental health of people affected by the disease. The diagnosis is pretty clear at this point but Ell now the means with which psychiatrists and medical practitioners could treat and monitor patients was not ideal. Their diagnosis would mostly depend on the patient's detailed insights into themselves. With a short and limited contact Time with each patient, there is a huge deficit in the data that must be needed by the medical practitioner for a solid diagnosis. That vacuum can be filled with data collected from smart watches or phones of people if it can be accurately translated to useful bits. The ease of accessibility makes digital phenotyping a very plausible global diagnosis system. With almost everyone already owning a smart phone, no additional device needs to be bought for diagnosis, thus producing nearly no additional costs to anyone.

2. LITERATURE REVIEW

Many scientists have always been interested in human activity recognition research. The development of a machine learning algorithm capable of identifying complicated human physical activity in real-world situations is still ongoing. There are several types of sensors that can be used utilized to identify the activity of humans Accelerometers are widely utilized since they come with a variety of functions which aid in the classification of activity

Researchers in [1] employed a smartphone's accelerometer to identify data instead of wearable sensors physical activity utilizing a machine learning method known as support vector machine and tracking it in real time. Time is a real person who uses an Android phone (OS). The accelerometer, gyroscope, and GPS data. The linearity of accelerometer sensors was accelerometer sensors was collected and pre-processed.

Researchers in [2] attempted to create a multi-layer parallel LSTM RNN model that could classify objects u sing smartphone sensor data to calculate human activity. Their studies yielded better outcomes than expected traditional machine learning approaches, as well as the outcomes provided by CNNs in general. Their LSTM model had a 94 percent accuracy rate. However, they discovered that LSTM RNNs were far less effective computationally more difficult than CNNs the network's computational complexity reduce with the number of LSTM units running concurrently. It was deduced from this that increasing the number of neurons in a parallel LSTM unit improved the network's performance marginally. The performance of the LSTM was improved by adding neurons to a parallel LSTM unit.

3. NEURAL NETWORKS

A neural network is a set of algorithms that aims to recognize underlying relationships in a batch of data using a method that replicates how the human brain works. In this context, neural networks refer to natural or artificial systems of neurons. They can discover hidden patterns and correlations in raw data using algorithms, cluster and categorize it, and learn and improve it over time.

3.1 Convolutional Neural Network

Convolutional Neural Networks are a type of neural network that is commonly used to recognize images. Its features such as the use of local receptive fields to identify specific aspects or features within the image input into it, which may subsequently be used to distinguish it from a set of similar images [6]. The pooling layer, non-linearity layer, convolutional layer, and fully - connected layers are all layers of a CNN. Convolution is a mathematical equation in which two components are merged to create a third function. The sensors used in our smart phones for measuring movement of the body is called Accelerometer and the one used to detect orientation is called Gyrometer. Their output, which is a time-series data set will be fed to the 1D CNN model. This model can map to different types of inputs and no manual input featuring is required.

3.2 Long Short-Term Memory

Out of the two RNN models, namely GRU which stands for Gated Recurrent Unit and LSTM we chose LSTM because it provides a higher accuracy and faster learning rate [4]. This too needs no manual input feeding system and is therefore, a good model to implement. LSTM networks are effortlessly served long data sequences to categorizing, reprocessing and fabricating prognostications predicated on time series data, since there can be lags of unknown time between vital episodes in a time series. LSTMs were evolved to deal with the vanishing gradient problem that can be faced when training conventional RNNs.

4. COMPARISON BETWEEN CCN AND LSTM NEURAL NETWORK

Convolutional Neural Network	Long Short-Term Memory
Consists of various layers that help in the processing of data and spatial correlation	Consists of various gates that help in making predictions about the data
Shows average performance with large datasets	Useful for processing large datasets
Does not retain data memory	Retains previous data unit

5. SOFTWARE USED IN HUMAN ACTIVITY RECOGNITION SYSTEM

Different supervised, semi-supervised and unsupervised algorithms can be used to solve the problem of real-time recognition. Different algorithms had proved themselves useful in different applications, which is why there is no clear distinction on which algorithm is more appropriate than others.

5.1 Machine Learning Algorithms

Based on the Supervised Learning technique, this is one of the simplest Machine Learning algorithms. The K-NN algorithm assumes that the new case/data and preexisting cases are comparable and places the new case in the category that is most similar to the existing categories.

The K-NN technique stores all available data and classifies a new data point based on its similarity to the existing data. This means that as fresh data is generated, the K-NN algorithm can quickly classify it into a suitable category [3]. The K-NN algorithm can be used for both regression and classification, but it is primarily employed for classification problems. The K-NN algorithm is a quasi-algorithm, which means it makes no assumptions regarding data. It's also known as a slow learner.

5.2 Support Vector or SVM

This is one of the most widely used Supervised Machine learning for Classification and Regression issues. However, it is mostly employed in Machine Learning for Categorization difficulties. The purpose of the SVM method is to find the optimum line or decision boundary for dividing n-dimensional space into classes so that additional data points can be readily placed in the correct category in the future.

5.3 Random Forest

This is a machine learning strategy used to address regression and classification problems. It employs ensemble learning, a technique that combines numerous classifiers to find solutions to complex problems. A random forest algorithm is made up of a lot of decision trees. The random forest algorithm's 'forest' is trained via bagging. Bagging is a meta-algorithm for improving the accuracy of machine learning systems.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 06 | Jun 2022 www.irjet.net p-ISSN: 2395-0072

Based on the predictions of the decision trees, the (random forest) algorithm determines the outcome. It forecasts by averaging or averaging the output from various trees. Increasing the number of trees improves the outcome's precision. It lowers the likelihood of overfitting.

6. OUTPUT FOR CCN AND LSTM

The programming code for this project was written in Python using the Spyder IDE. A sequential model was used for the implementation, which allows for the building of a model and then the addition of layers later. It simplifies the process of creating deep learning models. The model was created with a 3.0 GHz Hexa-Core AMD Ryzen CPU and 16 GB 3200 MHz LPDDR3 RAM on the online Kaggle platform IDE. The model was trained for 42 epochs with a batch size of 32. The data is first tested on the Sphere Dataset and then validated by the Validation Dataset called the WISDM Dataset.

6.1 Confusion Matrix- CNN Sphere Dataset

The CNN model achieves an accuracy of up to 77 percent on the testing dataset. The forecasts' absolute and normalized values are displayed in the confusion matrix. Despite the fact that the classifier's accuracy is around 77%, the overall performance cannot be considered satisfactory. There is a big difference in the results of each individual activity. Standing activity is the most accurate, with an accuracy of 83 percent. When it came to walking upstairs and downstairs, the classifiers performed horribly.

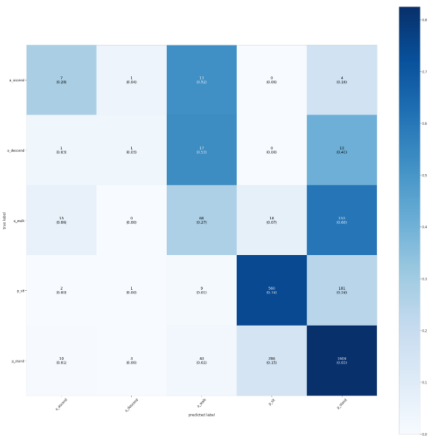

Fig 1. Confusion Matrix of Sphere Dataset for CNN

6.2 Confusion Matrix- CNN WISDM Dataset

When shifting from the testing set to the validation set, the consistency of the classifier's performance is crucial. On the WISDM dataset, the classifier has a rate of 21.5 percent accuracy. CNN can only track seated behavior incredibly effectively, according to the graph above, with a 99 percent accuracy rate. Walking, walking upstairs, and walking downstairs are all labelled as 'standing,' probably because it can't distinguish the difference. The negative output to be highlighted here is that it wrongly classifies standing activity as sitting, which was not the case with the testing set.

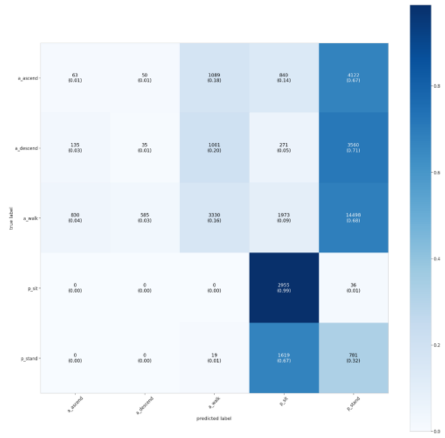

Fig 2. Confusion Matrix of WISDM Dataset for CNN

6.3 Confusion Matrix- LSTM Sphere Dataset

When the LSTM RNN was compiled numerous times, it had a 69 percent accuracy. The findings can be evaluated for each individual activity using the confusion matrix given in the figure above. The LSTM model follows the same pattern as the CNN model in that it does a good job of classifying stationary activities like sitting and standing but suffers with activities that need a lot of movement, including walking, walking upstairs, and walking downhill.

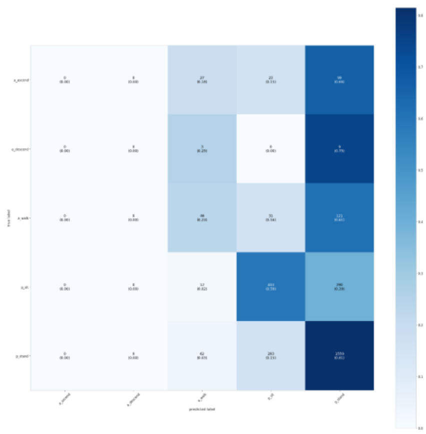

Fig 3. Confusion Matrix of Sphere Dataset for LSTM

6.4 Confusion Matrix- LSTM WISDM Dataset

The LSTM model outperformed the CNN model on the validation dataset, with an accuracy of 48 percent on several runs, in contrast to the CNN model's performance. Its performance on individual activities in the testing dataset is not significantly different from its performance on them. Walking accuracy increased from 23 percent on the testing set to 70 percent on the validation set, while sitting accuracy increased from 59 percent to 79 percent. However, standing accuracy declined from 81 percent to 21 percent.

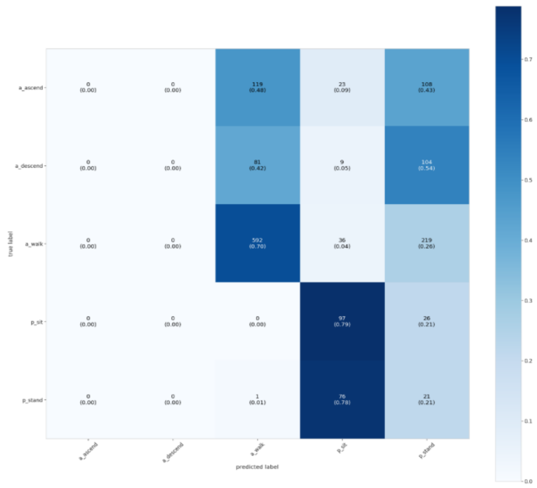

Fig 4. Confusion Matrix of WISDM Dataset for LSTM

7. COMPARISON BETWEEN OUTPUTS OF CCN AND LSTM NEURAL NETWORK

Parameters	CNN	LSTM
Learning Time	5:04:21 minutes	13:22:90 minutes
Retention of Memory (at 21 epochs)	CPU: 135% GPU: 20% RAM: 2.2GB GPU: 15.6GB	CPU: 108% GPU: 67% ZAM: 2GB GPU: 15.6GB
Retention of Memory (at 42 epochs)	CPU: 134% GPU: 25% RAM: 2.2GB GPU: 15.6GB	CPU: 102% GPU: 69% RAM: 2GB GPU: 15.6GB

Table 2. Execution Comparison of CNN and LSTM

Parameters	CNN	LSTM
Standing	83%	81%
Sitting	74%	59%
Walking	27%	23%
Walking Downstairs	3%	0%
Walking Upstairs	28%	0%
Overall Accuracy	77%	69%

Table 3. Confusion Matrix Comparison for Sphere Dataset of CNN and LSTM

Parameters	CNN	LSTM
Standing	32%	48%
Sitting	99%	79%
Walking	16%	70%
Walking Downstairs	1%	0%
Walking Upstairs	1%	0%
Overall Accuracy	21%	48%

Table 4. Confusion Matrix Comparison for WISDM Dataset of CNN and LSTM

8. LIMITATIONS

When it came to walking upstairs and downstairs, the classifiers performed horribly. There might be a number of reasons for these erroneous categorizations. One interesting feature is that activities like standing and sitting, which were identified more accurately than the rest in both our original and testing datasets, were more densely present than activities like walking, going upstairs, and walking downstairs. In the testing set, there were around 2000 instances of standing, but only 42 samples of walking upstairs. Moreover, due to the retention of memory in LSTM, the model can experience overfitting [7]. In this, the model gets very well acquainted to the program, resulting in absorbing noise and errors.

9. APLLICATIONS

The Human Activity Recognition System has vivid applications in many fields all over the world. This wide range of usage is due to high flexibility and correlation with the systems it is applied to alongside. This application can be further converted into a software application used for nursing, supervision and fitness.

9.1 Active living applications for smart homes

Modern technological advancements have created novel approaches to improve the quality of independent living for the elderly and disabled. A smart house, is an environment with sensors that improve occupants' safety and monitor their health problems. As a result, smart homes increase the freedom and quality of life of those who require assistance with physical and cognitive tasks. In general, the behavior of occupants and their interactions with the environment are monitored inside a smart home by evaluating data acquired through sensors.

9.2 Healthcare monitoring applications

The advancement of medical research and technology has significantly improved the quality of life for patients. According to Goldstone, life expectancy rates will rise considerably by 2050, with over 30% of Americans, Canadians, Chinese, and Europeans reaching the age of 60 [5]. As a result, researchers are attempting to improve existing healthcare monitoring systems in order to tackle urgent medical emergencies and reduce a patient's hospital stay and frequent medical visits.

9.3 Security and surveillance applications

Human operators supervise traditional surveillance systems. They should always be mindful of the human actions that are visible through the camera images. As the number of camera installations and views grows, the operators' task becomes more stressful, and their productivity suffers as a result. As a result, security companies are turning to vision-based technology for assistance in automating human operator operations and detecting anomalies in camera images.

10. FUTURE WORK

This paper can be used as a reference before building a human activity recognition system by applying either of the neural networks. An application can be built for online gymnasiums, surveillance and monitoring systems. A software for health tracking can be made for people with a smart phone, which can help doctors to get real-time data for treatment rather than the data manually filled in by the patient, which can be faulty. The Human Activity Recognition System has a wide range of uses in the security and surveillance sector as well. The Human Activity Recognition System could be a vital application needed in the future with all activities turning digital.

10. CONCLUSION

All models were unable to distinguish going upstairs and downstairs in both datasets – testing and validation – after comparing all parameters of CNN and LSTM. The findings observed might be due to a variety of factors. The LSTM RNN model has a greater accuracy on the validation dataset than the CNN model. This supports the theoretical assumption that LSTM RNNs are better at categorizing time series data than CNNs, which was established before these models were developed. This might also explain why the LSTM model fails to categorize actions such as walking upstairs and downstairs.

ACKNOWLEDGEMENT

We would like to express our gratitude to our professor, Ms. Uttara Bhatt and our parents for their guidance and immense support.

REFERENCES

[1] D. N. T. a. D. D. Phan, “Human Activities Recognition in Android Smartphone Using Support Vector Machine,” in 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), Bangkok, 2016.

[2] J. C. N. Y. a. X. L. T. Yu, “A Multi-Layer Parallel LSTM Network for Human Activity Recognition with Smartphone Sensors,” in 10th International Conference on Wireless Communications and Signal Processing, Hangzhou, 2018.

[3]Jhinal Modi, Prof. Hetal Bhaidasna, Jubin Bhaidasna, "Human Activity Recognition using Deep Learning Methods”

[4] Tahmina Zebin, Matthew Sperrin, Niels Peek and Alexander J. Casson, Senior Member, IEEE "Human activity recognition from inertial sensor time-series using batch normalized deep LSTM recurrent networks “

[5] J. A. Knight, “Physical Inactivity: Associated Diseases and Disorders,” 2012. [Online].

http://www.annclinlabsci.org/content/42/3/320.full
[Accessed 2nd October 2019]

[6] T. A. M. a. S. A.-Z. S. Albawi, “Understanding of a convolutional neural network,” 2017. [Online]. Available: https://ieeexplore.ieee.org/document/8308186
[Accessed 4 April 2020].

[7] N. e. a. Srivastava, “Dropout: a simple way to prevent neural networks from overfitting,” 2014. [Online]. Available:
https://jmlr.org/papers/v15/srivastava14a.html
[Accessed 11 April 2020].