IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Human Activity Recognition

A Review

[1]Saniya, [2]Falguni Rathiya, [3]Vaishnavi Wakde

Department of Computer Engineering, Dhole Patil College of Engineering, Wagholi, Pune, India

[4]Prof. Chaitanya Mankar, Assistant Professor

Department of Computer Engineering, Dhole Patil College of Engineering, Wagholi, Pune, India

Abstract: Nowadays human activity detection has come to be an implicit service in lots of smartphones. Everyone is turning into fitness conscious. Smartphones have embedded sensors consisting of accelerometers and gyroscopes. Therefore, the algorithms for such activity recognition should be very light to function on most smartphones or any wearables however correct at the identical time. During this paper, we will develop a lightweight software for human activity detection supported Long Short-Term Memory networks, which can learn features from raw accelerometer information, absolutely bypassing the process of generating manufactured features. We are going to assess our algorithm on information collected in a managed setting, in addition to information collected under field conditions, and we display that our algorithm is robust and performs almost equally well for each scenario even surpassing other approaches from the literature.

***Index Terms* - Accelerometer, Android Application, CNN, Gyroscope, Human Activity recognition, LSTM, Smartphones, SVM.**

I. INTRODUCTION

Human activity recognition (HAR) is a vast area of study mainly concerned with identifying specific movements or actions of a person based on given input data. The activities can be typical ones, like walking, sitting, running, walking on the staircase, sleeping, or more focused ones like eating, reading, or industrial operations. With the advancement of smart devices, wireless communication, and deep learning, daily human activity classification and recognition has got an increasing interest in areas like context-aware computing, smart assistive technologies for industries, where manual work remains dominant, and in rehabilitation centres where human motion monitoring is essential. Smartphones have embedded sensors such as in them like accelerometer, gyroscope, etc. in devices to enhance their usability, controllability, and management.

Traditional way to measure activity is by attaching special hardware devices on predefined location, like hip and ankle [1]. Sensor measurements from those devices are recorded on internal memory, to be later analyzed for different purposes [3]. Many of the epidemiological and clinical studies still use this method in their research [2]. With the technology improvement, body sensor networks allow more advanced approach, where sensor measurements can be sent directly to the users' smartphone to be analyzed on the fly [4]. In the last few years, modern smartphones are equipped with dozens of different sensors, therefore, smartphone measurements can be used for the process of activity detection, bypassing the need for extra hardware devices [5].

Hardware-friendly approach in [6] adapts the standard Support Vector Machine (SVM) to reduce computational cost while maintaining accuracy comparable to other traditional SVM based classification methods. More recent approaches are focused on features extraction from the raw acceleration data [1]. In [7], an unsupervised classification method is used for activity recognition [1].

The main objective of this paper is to develop a new lightweight algorithm for activity recognition, with the following characteristics: (i) to be easily implementable on mobile applications; (ii) to surpass other approaches from the literature using accuracy; and (iii) to be robust enough to perform almost equally good on data collected under field conditions as on data collected in a very controlled environment. We will train the model to acknowledge more activities. HAR has applications in healthcare, monitoring, and user recognition. We are developing the application for healthcare where we are counting footsteps on the activity recognized.

II. RELATED WORK

Although there are many techniques within the literature for activity recognition, during this paper we investigated a deep learning approach that supported Long Short-Term Memory (LSTM) networks and Convolutional Neural Network (CNN, or ConvNet).

LSTM:
[LSTMs are a special quite RNN, capable of learning long-term dependencies which make RNN smart at remembering things that have happened within the past and finding patterns across time to create its next guesses make sense. LSTMs broke records for improved Machine Translation, Language Modelling, and Multilingual Language Processing.[8]]

Advantages:

- [Neural Networks is a very powerful technique and is used for image recognition and many other applications.
- RNN addresses that issue by including a feedback look which serves as a kind of memory. So, the past inputs to the model leave a footprint. LSTM extends that idea and by creating both a short-term and a long-term memory component.
- Hence, LSTM is great tool for anything that has a sequence. Since the meaning of a word depends on the ones that preceded it. This paved the way for NLP and narrative analysis to leverage Neural Networks.
- LSTM can be used for text generation. You can train the model on the text of a writer, say, and the model will be able to generate new sentences that mimics the style and interests of the writer.
- Sequence-to-Sequence LSTM models are the state of the technique for translations. They also have a wide array of applications like time series forecasting.[13]]

Disadvantages:
[As it is said, everything in this world comes with its own advantages and disadvantages, LSTMs too, have a few drawbacks which are discussed as below:

- LSTMs became popular because they could solve the problem of vanishing gradients. But it turns out, they fail to remove it completely. The problem lies in the fact that the data still has to move from cell to cell for its evaluation. Moreover, the cell has become quite complex now with the additional features (such as forget gates) being brought into the picture.
- They require a lot of resources and time to get trained and become ready for real-world applications. In technical terms, they need high memory-bandwidth because of linear layers present in each cell which the system usually fails to provide for. Thus, hardware-wise, LSTMs become quite inefficient.
- With the rise of data mining, developers are looking for a model that can remember past information for a longer time than LSTMs. The source of inspiration for such kind of model is the human habit of dividing a given piece of information into small parts for easy remembrance.
- LSTMs get affected by different random weight initializations and hence behave quite similar to that of a feed-forward neural net. They prefer small weight initializations instead.
- LSTMs are prone to overfitting and it is difficult to apply the dropout algorithm to curb this issue. Dropout is a regularization method where input and recurrent connections to LSTM units are probabilistically excluded from activation and weight updates while training a network.[12]]

CNN:
[Next comes the Convolutional Neural Network (CNN, or ConvNet) which is a class of deep neural networks that are most commonly applied to analyzing visual imagery. Their other applications include video understanding, speech recognition, and understanding natural language processing.[8]]

Advantages:

- [The usage of CNNs are motivated by the fact that they can capture / are able to learn relevant features from an image /video at different levels similar to a human brain. This is feature learning! Conventional neural networks cannot do this.
- Another main feature of CNNs is weight sharing. Let's take an example to explain this. Say you have a one layered CNN with 10 filters of size 5x5. Now you can simply calculate parameters of such a CNN, it would be 5*5*10 weights and 10 biases i.e., 5* 5*10 + 10 = 260 parameters. Now let's take a simple one layered NN with 250 neurons, here the number of weight parameters depending on the size of images is '250 x K' where size of the image is P X M and K = (P *M). Additionally, you need 'M' biases. For the MNIST data as input to such a NN we will have (250*784+1 = 19601) parameters. Clearly, CNN is more efficient in terms of memory and complexity. Imagine NNs and CNNs with billions of neurons, then CNNs would be less complex and saves memory compared to the NN.
- In terms of performance, CNNs outperform NNs on conventional image recognition tasks and many other tasks. Look at the Inception model, Resnet50 and many others for instance.
- For a completely new task / problem CNNs are very good feature extractors. This means that you can extract useful attributes from an already trained CNN with its trained weights by feeding your data on each level and tune the CNN a bit for the specific task. E.g.: Add a classifier after the last layer with labels specific to the task. This is also called pre-training and CNNs are very efficient in such tasks compared to NNs. Another advantage of this pre-training is we avoid training of CNN and save memory, time. The only thing you have to train is the classifier at the end for your labels.[14]]

Disadvantages:

- [A Convolutional neural network is significantly slower due to an operation such as maxpool.
- If the CNN has several layers then the training process takes a lot of time if the computer doesn't consist of a good GPU.
- A ConvNet requires a large Dataset to process and train the neural network.[15]]

[Also, LSTM combined with Convolutional Neural Networks (CNNs) improved automatic image captioning like those are seen in Facebook. Thus, you can see that RNN is more like helping us in data processing predicting our next step whereas CNN helps us in visuals analysis.

Though RNNs operate over sequences of vectors: sequences in the input, the output, or in the most general case both in comparison with CNN which not only have constrained Application Programming Interface (API) but also fixed amount of computational steps. This is why CNN is kind of more powerful now than RNN. This is mostly because RNN has gradient vanishing and exploding problems (over 3 layers, the performance may drop) whereas CNN can be stacked into a very deep model, for which it's been proven quite effective.

But CNNs are not also flawless. A typical CNN can tell the type of an object but can't specify their location. This is because CNN can regress one object at a time thus when multiple objects remain in the same visual field then the CNN bounding box regression cannot work well due to interference.[8]]

When using sensor data to train a model, the accuracy is generally high if the training set and test set belong to the same collection of users. When it comes to different user sets, the accuracy rate will drop obviously. We've built an LSTM model that can predict human activity from 200 time-step sequence with over 97% accuracy on the test set [10]. The resultant data was given to the designed CNN (Convolutional Neural Network) for classification. Both data structuring methods were analyzed and compared yet the time series data structuring showed a better result and attained an accuracy of 99.5% [11].
we will be implementing our system in an android application that will continuously recognize current activity being performed and calculates metrics such as step count and sets the phone modes accordingly.

III. CONCLUSION

In this paper, we have discussed human behaviors through their actions and recognized corresponding activities. As we discussed, Human activity recognition has broad applications in the human survey system and medical research. In this project, we have designed a smartphone-based recognition system that recognizes six human activities: walking, sitting, standing, lying, going upstairs, and going downstairs. The system collected time-series signals using a built-in accelerometer and gyroscope functionalities. The activity data were trained and tested using 2 machine learning methods: Long short-term memory networks (LSTM) and convolutional neural networks (CNN) in artificial intelligence.

REFERENCES

[1] M. Milenkoski, K. Trivodaliev, S. Kalajdziski, M. Jovanov and B. R. Stojkoska, "Real time human activity recognition on smartphones using LSTM networks," 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2018, pp. 1126-1131, doi: 10.23919/MIPRO.2018.8400205.
[2] Zdravevski, Eftim, Biljana Risteska Stojkoska, Marie Standl, and Holger Schulz. "Automatic machine-learning based identification of jogging periods from accelerometer measurements of adolescents under field conditions." PloS one 12, no. 9 (2017): e0184216.
[3] Robusto KM, Trost SG. Comparison of three generations of ActiGraph™ activity monitors in children and adolescents. Journal of Sports Sciences. 2012;30(13):1429–1435.
[4] Fortino, Giancarlo, Stefano Galzarano, Raffaele Gravina, and Wenfeng Li. "A framework for collaborative computing and multi-sensor data fusion in body sensor networks." Information Fusion 22 (2015): 50-70.
[5] Su, Xing, Hanghang Tong, and Ping Ji. "Activity recognition with smartphone sensors." Tsinghua Science and Technology19, no. 3 (2014): 235-249.
[6] Anguita, Davide, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge L. Reyes-Ortiz. "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine." In International workshop on ambient assisted living, pp. 216-223. Springer, Berlin, Heidelberg, 2012.
[7] Lu, Yonggang, Ye Wei, Li Liu, Jun Zhong, Letian Sun, and Ye Liu. "Towards unsupervised physical activity recognition using smartphone accelerometers." Multimedia Tools and Applications 76, no. 8 (2017): 10701-10719.
[8] https://medium.com/@sprhlabs/understanding-deep-learning-dnn-rnn-lstm-cnn-and-r-cnn-6602ed94dbff
[9]https://medium.com/@curiousily/human-activity-recognition-using-lstms-on-android-tensorflow-for-hackers-part-vi-492da5adef64
[10] T. T. Alemayoh, J. Hoon Lee and S. Okamoto, "Deep Learning Based Real-time Daily Human Activity Recognition and Its Implementation in a Smartphone," 2019 16th International Conference on Ubiquitous Robots (UR), Jeju, Korea (South), 2019, pp. 179-182, doi: 10.1109/URAI.2019.8768791.
[11] https://www.geeksforgeeks.org/understanding-of-lstm-networks
[12] https://www.quora.com/What-are-the-advantages-of-LSTM-in-general
[13]https://www.quora.com/What-are-the-advantages-of-a-convolutional-neural-network-CNN-compared-to-a-simple-neural-network-from-the-theoretical-and-practical-perspective
[14] https://iq.opengenus.org/disadvantages-of-cnn/