

Human Activity Recognition Based on Smartphone Sensor Data Using CNN

Kenan İsmail *, Kasım Özacar

Department of Computer Engineering, Karabuk University, Karabuk, Turkey

kenanheso@gmail.com, kasimozacar@karabuk.edu.tr

KEYWORDS: CNN, Sensing data, human activity recognition

ABSTRACT:

Human activity recognitions have been widely used nowadays by end users thanks to extensive usage of smartphones. Smartphones, by self-containing low-cost sensing technology, can track our daily activities for serving healthcare, sport, interactive AR/VR games and so on. However, smartphone technology is evolving and the techniques of using the data that smartphones go through are also improving. In this study, we used built-in sensing technologies (accelerometer and gyroscope) available in nearly every smartphone to detect the most common 5 daily activities of human by taking the data of these sensors and extract the features for a Convolutional Neural Network (CNN) model. We prepare a dataset and use TensorFlow to train the collected data from the sensors then filtered it to be processed. We also discuss the differences in CNN model accuracy with different optimizers. To demonstrate the model, we developed an android application that successfully predict an activity. We believe that after improving this application, it can be used for especially lonely old people to immediately warn authorities in case of any daily incidents.

1. INTRODUCTION

In recent years, smartphones have been running rapidly in parallel with our life, all smartphones have some self-contained sensors, some of these sensors can record large amounts of data, many smartphone applications use this data for multiple purposes such as counting the number of steps, measuring the heart rate, heartbeat, or they can be used in the fitness field and so on. However, smartphone technology is evolving and the techniques of using the data that smartphones go through are also improving (Lara et al., 2012, Shoaib M et al., 2016, Majumder et al 2019). There exists plenty of feature extraction techniques classifying sensory data such as ANN, KNN and SVM (Ronao, C.A et al., 2014, Seera, M. et al., 2014, Eastwood, M et al., 2014). Unlike these conventional approaches, deep neural network learns features directly from the input data without requiring manual feature extraction operation.

In this study, we have utilized two built-in sensors (3 axis acceleration and 3 axis gyro sensor data) available inside nearly in every smartphone to detect the daily activities of human by taking the data of these sensors and extract the features using Convolutional Neural Network (CNN) to recognize human activities. We show how to collect data from embedded sensors in smartphones then how to filter the data to be processed and explain the strategy which we used to shape the dataset before and after processing it with TensorFlow and Keras libraries. Also, we show the differences in CNN model accuracy with different optimizers as well as explain the different layers of CNN model. Our implemented system works efficiently in self-contained smartphones and provides adequate activity recognition (95.83%).

2. DATA COLLECTION

We collect data using Samsung Galaxy S7 android device, and record data in 100Hz. The dataset contains 9 columns which are user id, timestamp, activity name and the data of 6 axis accelerometer and gyroscope (ax, ay, az, gx, gy, gz) data. To prepare dataset, we develop an android application in which the sensor data has been saved to a text file. We put the smartphone in pocket in different positions for each data recording. We choose two sensor data which are available in almost every smartphone: Linear Acceleration and Rotation Vector. Linear acceleration sensor measures the acceleration force in m/s^2 on 3D axes (X, Y and Z) excluding the force of gravity while the Accelerometer sensor include this force. A rotation vector measures the orientation of the device by providing the three elements of device's rotation vector.

3. DATA PROCESSING

3.1 Load data and balance of activity classes

The dataset contains 9 columns for each row and gathered in one text file. We used NumPy and Pandas libraries to state and process the dataset. Then, we defined 5 activities to classify that are (Walking, Sitting, Standing, go upstairs, go downstairs). Since activity distribution must have the same number of samples, we balance these activities by simply making each activity has the same number of data samples.

3.2 Standardizing data with StandartScaler

Since 6-axis sensor data does not have the same variance, we bring them into the same variance level by applying standardization. The standard score of a sample x is

*Corresponding author

calculated as: $score = (x - u) / s$, where u is the mean, and s is the standard deviation of the training samples.

3.2 Data Segmentations

We choose to have one window of data to represent 4 seconds of an activity. We created two-dimensional data frames with 100 Hz for a single activity recognition, resulting 2400 (6 sensor data x 400) window size. We randomly pick 80 % of data for training and 20% for testing.

4. DEEP NEURAL NETWORK MODEL

A CNN (Convolutional Neural Network) model, one of the most famous deep neural network models, has many layers that are used mainly for auto correlated data or image processing. The conv layer is the main layer in CNN, it varies between 1D, 2D and 3D. In our project, we chose Conv2D which has given an accuracy of 95.83%. First, we identified a sequential model that later we added other layers too. Then we added 5 more layers. While the first Conv2D layer creates a convolution kernel that is convolved with the layer input to produce a tensor of outputs, the second layer is also a Conv2D layer with increasing filter size. Then we added the flatten layer which is a key step in all Convolutional Neural Networks and has no effect on the batch size, and after that we added a dense layer, in which all inputs and outputs are interconnected and each has a separate weight, with a filter value of 64 and a ReLU(Rectified Linear unit) activation function. The last layer is also a Dense layer with Softmax activation function and has 5 units to represent the number of classes. Then we compiled with RMSprop and Adam optimizers to find out which one is the best in terms of accuracy.

Details of the model are given in Table 1.

Layer	Parameters	Activation
Conv2D	16 filter size, input shape	ReLU
Conv2D	32 filter size	ReLU
Flatten		
Dense	64 filter size	ReLU
Dense	Num_classes	softmax

Table 1. Layer in Deep Neural network

5. ANDROID APPLICATION

Figure 1 shows the application we developed to demonstrate how the model works in smartphones. However, trained CNN model has been saved in TensorFlow 2.2.1 version. Therefore, to perform prediction in a mobile device we converted CNN model to

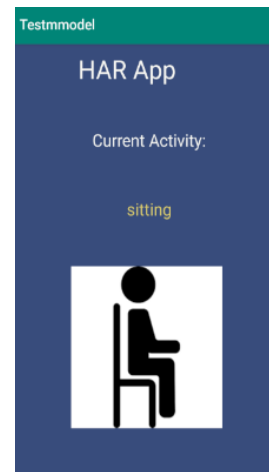


Figure 1. Android application to recognize activities.

TensorFlow lite version and imported the necessary libraries to TensorFlow lite model inside it later. The application simply listens user activity for 4 seconds and both displays and loudly localizes it.

6. RESULTS

We discussed the obtained results through the study, since we compiled the CNN model with different optimizers

Optimizer	Model evaluation
RMSprop	95.83%
Adam	83.33%

Table 2. Accuracy respect to optimizer.

(RMSprop and Adam). As can be seen in Table 2, RMSprop gives better prediction than Adam optimizer. We also filtered the raw data to reduce amount of variance in neural network model.

7. CONCLUSION

Using smartphone leads a new way of interaction thanks to its low-cost self-containing sensing technology. With these sensors, they have been used for healthcare, AR/VR interaction, sports and so many purposes. Therefore, in this study, we utilized a smartphone to track daily activities for serving healthcare. We used built-in sensing technologies (accelerometer and gyroscope) in smartphones to detect the daily activities of human by taking the data of these sensors and extract the features for a Convolutional Neural Network (CNN) model. We prepare a dataset and use TensorFlow to train filtered dataset and compare the accuracy of models for different optimizers. We finally developed an android app to demonstrate our model. We believe that this app can be improved to track and classify activity of especially lonely

old people to immediately warn authorities regarding their activity in case of any daily incidents.

REFERENCES

- Lara, O. D. , & Labrador, M. A. (2012). A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys & Tutorials*, 15 (3), 1192–1209.
- Majumder, Sumit, and M. Jamal Deen. "Smartphone sensors for health monitoring and diagnosis." *Sensors* 19.9 (2019): 2164.
- Team, K. (n.d.). Keras documentation: Keras layers API. Retrieved June 18, 2020, from <https://keras.io/api/layers/>
- Shoaib M., Bosch S., Incel O.D., Scholten H., Having a P.J.M. Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors*. 2016;16:426. doi: 10.3390/s16040426.
- Ronao, C.A.; Cho, S.-B. Human activity recognition using smartphone sensors with two-stage continuous hidden Markov models. In *Proceedings of the 2014 10th International Conference on Natural Computation (ICNC)*, Xiamen, China, 19–21 August 2014; pp. 681–686.
- Seera, M.; Loo, C.K.; Lim, C.P. A hybrid FMM-CART model for human activity recognition. In *Proceedings of the 2014 IEEE International Conference on Systems, Man and Cybernetics—SMC*, San Diego, CA, USA, 5–8 October 2014; pp. 182–187.
- Eastwood, M.; Jayne, C. Evaluation of hyperbox neural network learning for classification. *Neurocomputing* 2014, 133, 249–257.