Heterogeneous Human Activity Recognition with Dynamic Sensor Fusion using Deep Learning Models

Serlin Tamilselvam, Bhuvaneshwaran Ravi, Jayashree Srinivasan
Bishop's University, Sherbrooke, Quebec, Canada

Email: {stamilselvam19, bravi19, jsrinivasan19}@UBishops.ca

***Abstract*– Human activity recognition is one of the challenging concept for the real-world application. To recognize variety and wide range of activities performed by humans, is not a simple feat to achieve. To tackle this heterogeneity, we can use combination of sensors from smartphone and smartwatch devices. But sometimes users may deviate or do a different activity which poses difficulty to existing recognition algorithms. To minimize such inconsistency and to boost accuracy, we have implemented deep learning concepts and developed two models with Feed Forward Fully Connected Network and Recurrent Neural Network adaptive frameworks. We fuse the inputs from two sensors – accelerometer and gyroscope of smartphone and smartwatch, which is in turn feed in to train the models. Later we used cross visualization technique to evaluate the efficiency of the models. The result demonstrates a practical implementation of our model to existing human recognition algorithms and its advantages over existing approach.**

Keywords — Human Activity Recognition, Deep Learning, Neural Network, Feed Forward Fully Connected Network, Recurrent Neural Network.

I. INTRODUCTION

Our current time poses a unique opportunity to explore the massive data obtained from mobile and smart watches, Understanding the human activity can lead to tremendous development in various fields such as medicine, transportation, security, human behavior, elderly care, abnormal activity detection and even could be of great tactical value to military services

The activity recognition is an age old classical problem, many machine learning algorithms are leveraged in the past to detect an activity like random forest classifier, naïve bayes classifier and knn algorithms. However all the mentioned classifiers are considered to be static learning, we have to predefine which type of activity are going to detect. This assumption maybe workable in few situation but not practical for many real-life activities like base-jumping, logging etc. Some of the abnormally discarded activity as 'meaning less' can be normal activity among others. So defining all the meaning full activities is practically impossible in real-world applications.

When a user label such activities based on the historical data and if there is a new similar activity, system won't be able to identify the new activity and misclassify it. For example, the heartrate sensor can detect both exercising and panic attack and classical machine learning algorithm will be predefined to identify only exercise target. To overcome such unforeseen activities, heterogeneous activity recognition is proposed were a wide range sensor are in play to model the neural network, in the foreseen scenario both the heartrate sensor and custom barometer.

To summarize, we develop a practical dynamic heterogeneous sensor fusion framework, which addresses the challenge of dynamic sensor fusion in adaptive activity recognition. The key contributions of the paper are highlighted as follows:

- We propose a sensor fusion framework to learn sensor weights for each activity class so that activities are easier to be discriminated in the new distance space. We implement several neural network models like FNN and RNN.
- In contrast to prior models, this framework will learn the activities without any prior definition of any target activities.
- Experimental results on the provided dataset are encouraging and tested the effectiveness of the proposed framework on the activity recognition task with fused sensor inputs.

The following content of this paper will be in the order as follows. We will start with the potential sensors which can be used, various activities that can be predicted. After that, we will discuss the models implemented with the UCI datasets, the results obtained from both models, and comparison between their accuracy. Finally we conclude the paper with summary and potential implementation of this framework onto different real-world applications.

II. DATA AND METHODOLOGY

A. *Sensors and data acquisition*

The datasets are provided by the UCI machine learning repository. Acquisition of that data is done from Smartphones and Smartwatches is a dataset devised to benchmark human activity recognition algorithms (classification, automatic data segmentation, sensor fusion, feature extraction, etc.) in real-world contexts; specifically, the dataset is gathered with a variety of different device models and use-scenarios, in order to reflect sensing heterogeneities to be expected in practical deployments. The working of sensors are as follows

An accelerometer is made up of multiple axes, two to determine most two-dimensional movement with the option of a third for 3D positioning. Most smartphones typically make use of three-axis models, whereas cars simply use only a two-axis to determine the moment of impact. The sensitivity of these devices is quite high as they're intended to measure even very minute shifts in acceleration. The more sensitive the accelerometer, the more easily it can measure acceleration. A gyroscope works on Angular velocity sensing, it works on a 3-axis acceleration sensor X, Y, Z. It consists of 3 double T structure elements which reacts to shifts and movements. Other potential sensors can be used in the proposed model are Magnetometer, GPS, Proximity sensor

B. *Features and Targets*

The features of the dataset are the sensor coordinates, timeframe of the instance, user identification and the device identification. The targets are primarily classified into 6 different activities – Biking, Siting, Standing, Walking, Stair Up and Stair Down

Biking - using accelerometer to calculate the speed of the user and comparing it to the speed of predefined walking vectors.

Sitting - the height of an average human being is matched with the altitude form the ground level.

Standing - sitting and standing both are predicted in the same way that is altitude calculation from current ground level.

Walking - consist of backward walking, side way walking and a wide range of activities that are for the foot sensors.

Stair up - using sensors, if the altitude keeps increasing on a given scale, it is calculated walking up the stairs.

Stair down - going up and down the stairs are predicted the same way. In upstairs it is increasing.

III. INPUT ADAPTATION AND NETWORK ARCHITECTURE

The input feed from the datasets are down sampled to equalize the target classes and then selected combination of features are fed into the model – X,Y,Z coordinates, user identification and target values from sensor.

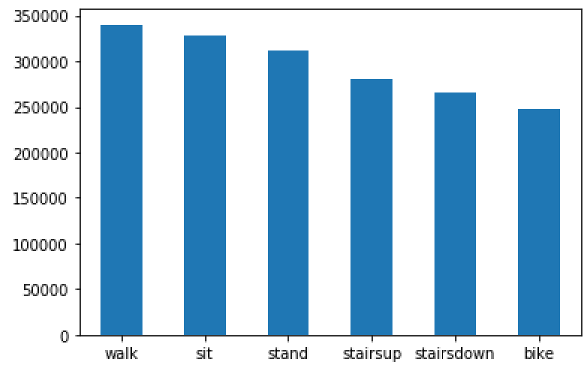

Fig. 1. The raw data that is fed into the framework.

A. *Feed Forward Neural Network Architecture*

We had initially used Neural Networks for classification but it did not make much sense to use just plain Neural Nets as the data which we were dealing with was time-series data. So we tried Convolution Neural Networks, but the network generally works best with multidimensional image dataset, so we settled with fully connected feed forward,

The network takes dynamic features as input which flows from the input layer to the hidden layers and finally flows through output layer. Each node is connected to every other node on the adjacent layers which is similar to how neurons are connected inside human brain. The optimal hyper parameters of the network are as follows

Number of hidden Layers – 4
Neurons – 100
Epochs – 0 to 50
Learning Rate – 0.001

The hidden layer contains linear function which takes input and multiple it with random weights $A = cx$ and it is activated by the ReLu activation function $A(x) = max(0,x)$, It gives an output x if x is positive and 0 otherwise.

Finally at the output layer we used softmax function were we obtained the probabilities for the arbitrary real values.

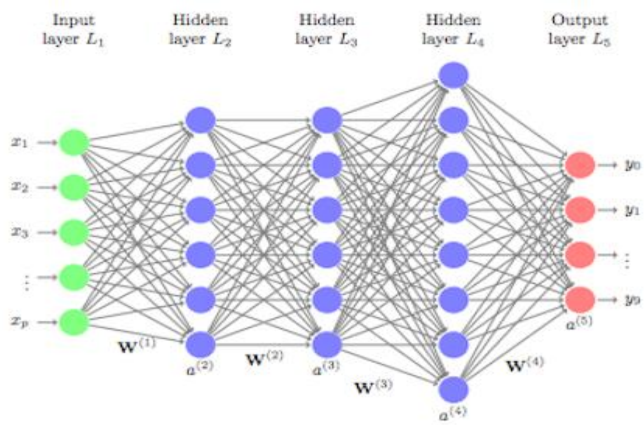

Fig. 2. Feed Forward Neural Network

IV. Results and Model Comparison

A Model Accuracy and Model Loss

Both the gyroscope as well as the accelerometer data gave extrapolated with accuracy up to 71%. The percentage accuracy on trained as well as test data are shown in the while tuning optimal hyper parameters. The best result was obtained for ReLu activation function with aforementioned hyper parameters

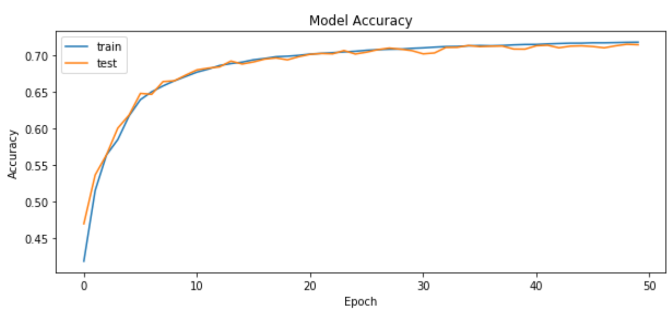

Fig. 3. Model Accuracy

We were able to reduce the model loss from 1.8% to 0.8% during the training process of the model.

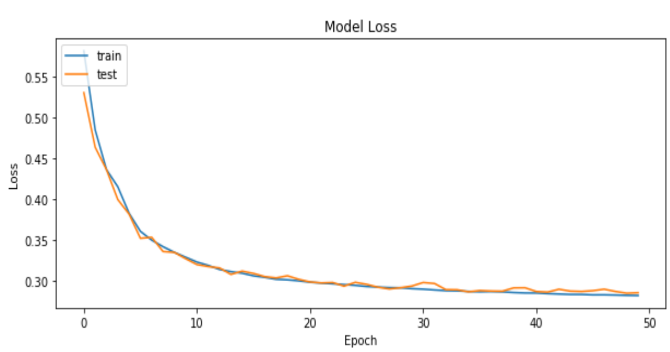

Fig. 4. Model Loss

For understanding the performance the proposed model, we designed a confusion matrix to see the model evaluation of the input data against the predicted data, the model gave ~80% prediction on four of the six targets fed into it.

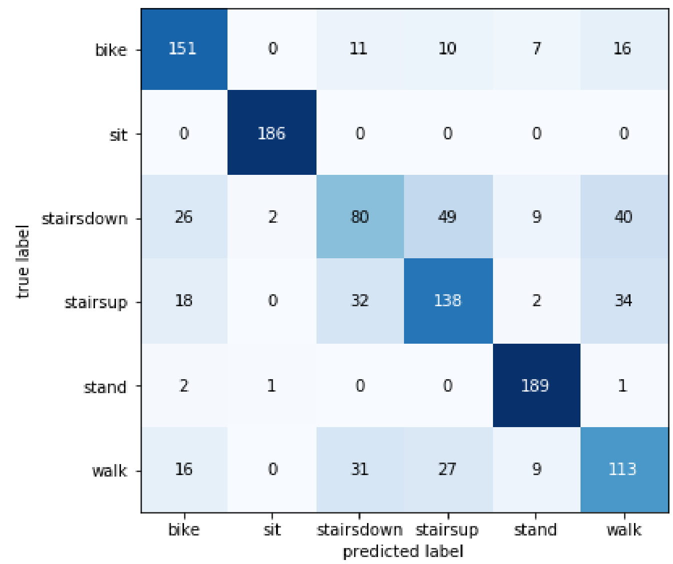

Fig. 5. Confusion Matrix for FNN

B. Comparison with Recurrent Neural Network model

We implemented Recurrent Neural Network to compare our model efficiency with other deep learning technique. The network proved difficult to train and achieve efficiency as the output of the previous time step is fed into the next time step. Through, the network is theoretically capable of higher prediction rate due to the capability to store data over the network cycle, it is extremely complicated to work with. The confusion matrix proves that the model was able to predict ~65% for two of six target classes.

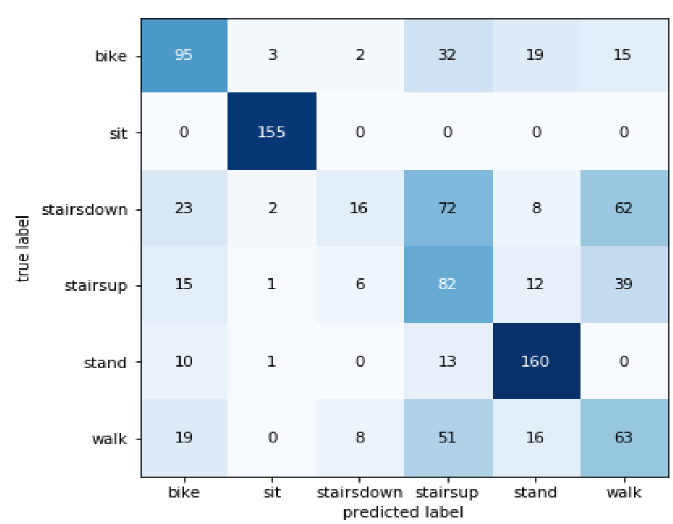

Fig. 6. Confusion Matrix for RNN

We were able to achieve 56.22% of model accuracy by tuning with the below hyper parameters.

Recurrent Cycle - 1
Neurons – 10
Epochs – 50
Learning Rate – 0.01

V. CONCLUSION

In this paper, we developed a sensor fusion framework for heterogeneous activity recognition using feedforward fully connected neural network. The framework can evaluate the importance of each class of sensor data for each activity captured and adjust the sensor weights incrementally. Based on the proposed framework, we can apply various real world application. The experiments were carried out on a Heterogeneity Activity Recognition Data Set from UCI repository which was collected using smartphone and smartwatch and promising results were obtained. As future work, we will assess our framework on more datasets and examine its efficiency of recognizing more complicated datasets

VI. CONTRIBUTIONS

- Research on all suggested projects - Jayashree Srinivasan
- Research on scientific papers – Bhuvaneshwaran Ravi, Serlin Tamilselvam
- Data Processing – Serlin Tamilselvam
- Convolutional Neural Network – Serlin Tamilselvam
- Feedforward Neural Network - ,Jayashree Srinivasan, Bhuvaneshwaran Ravi
- Recurrent Neural Network – Bhuvaneshwaran Ravi, Serlin Tamilselvam
- Report preparation – Bhuvaneshwaran Ravi, Jayashree Srinivasan

VII. REFERENCES

- Heterogeneity Activity Recognition Data Set from UCI repository, https://archive.ics.uci.edu/ml/datasets/Heterogeneity+Activity+Recognition
- Ming Zeng, Xiao Wang, Le T. Nguyen, Pang Wu, Ole J. Mengshoel, Joy Zhang, Adaptive Activity Recognition with Dynamic Heterogeneous Sensor Fusion
- Antonio Bevilacqua, Kyle MacDonald, Aamina Rangarej, Venessa Widjaya, Brian Caulfield, and Tahar Kechadi, Human Activity Recognition with Convolutional Neural Networks
- Minh-Son Daoa,b,*, Tuan-Anh Nguyen-Giab , Van-Cuong Maib, Daily Human Activities Recognition Using Heterogeneous Sensors from Smartphones