

Human Activity Recognition with Smartphones

Zhou Fang

Oregon State University
Corvallis, Oregon, U.S.A
fangzho@oregonstate.edu

Jiale Liu

Oregon State University
Corvallis, Oregon, U.S.A
liujial@oregonstate.edu

Abstract

Human activity recognition is an important technology which can be applied in many real-life scenarios such as eldercare and healthcare. Accurate prediction of indoor activities gains more attention recently. There are different methods have already used on this area. In this paper, we mainly discuss two methods on HAR, MLP and LSTM. We also have improved methods on these two. Feature selection will gain a superior performance with less number of features. We examine the classification performance of our approaches on a publicly available HAR dataset. The superiority of our approaches is verified by comparing their performances to other related work on HAR. An overall classification accuracy of 95.83% can be achieved.

1. Introduction

Recognition of human daily physical activities has recently been an active topic. Nowadays, there are many researches can achieve recognition of human indoor activities with simple sensors rather than cameras recording and tracking their activities everyday [1] [2]. Recognition technology becomes more popular because it can understand the human behavior, particularly for disabled and elderly people rather than videos which will reveal privacy.

The use of smartphones for wearable sensing introduces a more convenient HAR alternative with worthy advantages, e.g., it is shown to be nonintrusive to people's life on daily basis [3]. It has been reported that smartphones have become a considerable part of everybody's daily life and 97% of the world population use mobile phones. So, the advantage of smartphones help promotes the adoption of smartphones for HAR applications.

2. Related work

Some of the pioneering works in HAR using the

accelerometer and gyroscope was published in 2017 (Kotaro Nakano, & Basabi Chakraborty, 2017). However, the overall performance of K-NN and CNN in that published paper is 90.12% and 90.56% separately. K-NN shows a good performance on WALKING, LAYING, and STANDING but low accuracy on predicting WALKING DOWNSTAIRS, WALKING UPSTAIRS and SITTING. CNN shows a good performance on WALKING, WALKING DOWNSTAIRS and WALKING UPSTAIRS, but poor predicting on SITTING, LAYING and STANDING. That paper also worked on MLP model, however, the overall accuracy of their MLP is only 90.14% [1].

3. Proposed approaches

We work on HAR using two different models. The first model is Multilayer Perceptron. The second one is Long-Short Term Memory of Recurrent Neural Network model.

3.1. MLP

For MLP model, our inputs are 561 numeral data from two sensors (accelerometer and gyroscope). The 561 by 1 matrix goes through one hidden layer with 512 units. The active function of the hidden layer is Relu function. Finally, we get a 6 by 1 matrix at output layer of six classes probabilities with soft max classifier as Figure 1. The label of maximum in output matrix is the predicted activity for this instance.

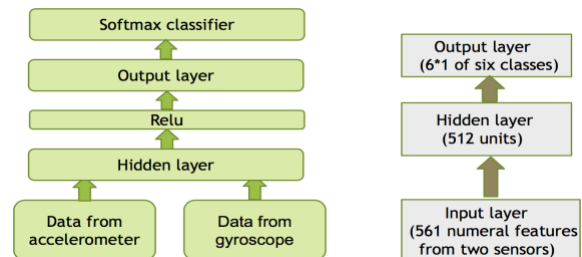


Figure 1. Structure of MLP Model

3.2. LSTM

For Long-Short Term Memory of Recurrent Neural Network, our inputs are 561 numeral data. We apply 512 recurrent units with LSTM cells in hidden layer. Similarly, we get six probabilities at the output layer as Figure 2. Then we pick the label of the maximum value as our prediction.

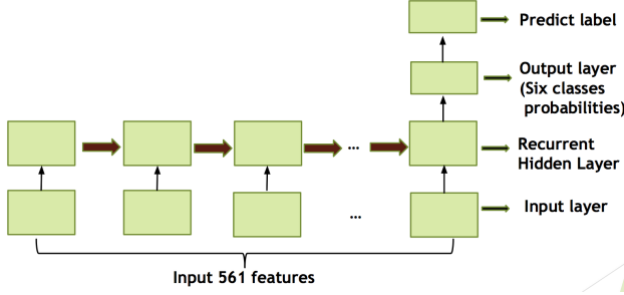


Figure 2. Structure of LSTM Model

4. Improvement

For improving our performance, we try to use more complicated model first, like adding more hidden layers, change optimizer, but none of these helps to our models. Then we propose to do feature engineering. Although using extracted 561 features can achieve higher recognition performance of different types of body movements, using all features directly requires a more complex model of long training and classification times.

In many real-time applications, a simple model with less computational cost and time is desired. We therefore proposal a feature selection step for the purpose of data reduction, with the redundant information being removed. Many feature selection methods have been reported in the literature. In this work, we investigate a feature selection based on importance from RandomForest, mutual information feature selection(MIFS) and F-score feature selection. Only feature selection based on importance from RandomForest works well on our model. Other two methods show worse performance than based case.

5. Simulation experiments

We discuss two parts in simulation experiments: HAR dataset, implementation of two proposed models.

5.1. Dataset

This data set is collected from 30 subjects within an age group of 19-48 years old. Each person performed a protocol of six activities (WALKING, WALKING UPSTAIRS, WALKING DOWNSTAIRS, SITTING, STANDING,

LAYING) wearing a smartphone (Samsung Galaxy SII) on the waist. The embedded accelerometer and gyroscope captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The experiments have been video-recorded to label the data manually. The obtained dataset has been randomly partitioned into two sets, where 70% of the volunteers was selected for generating the training data and 30% the test data. The sensor signals were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 reading windows). From each window, a vector of 561 features was obtained by calculating variables from the time and frequency domain. The details can be found in [4].

Conclusion, we get 561 numeral data in feature domain and 6 labels for predicting activities. All feature data have been normalized and within $[-1, 1]$.

5.2. Implement of simulation experiments

There are four parts, MLP and LSTM without feature selection, MLP and LSTM with feature selection based on importance from RandomForest(importance MLP and importance LTSM). We calculate the importance of each feature to final result by RandomForest. Only when the importance is greater than 0.1%. We select it as our feature, otherwise, drop it. In this way, 531 features are selected from 561 features.

All of our models implement on GPU, using library: Pytorch, Pandas, sikit-learn and skfeature from ASU.

Table 1 to Table 4 are the parameter setting of four models.

Items	Explanation of items
Hidden Dimension	512
Mini Batch Size	32
Optimizer	SGD (lr=0.0009, momentum=0.9)
Criterion Function	Cross Entropy
Num. of Features	561

Table 1. Parameter Setting of MLP

Items	Explanation of items
Hidden Dimension	512
Mini Batch Size	32
Dropout	0.5
Optimizer	SGD (lr=0.0009, momentum=0.9)
Criterion Function	Cross Entropy
Num. of Features	561

Table 2. Parameter Setting of LSTM

Items	Explanation of items
Hidden Dimension	512
Mini Batch Size	32
Optimizer	SGD (lr=0.0009, momentum=0.9)
Criterion Function	Cross Entropy
Num. of Features	531

Table 3. Parameter Setting of Importance MLP

Items	Explanation of items
Hidden Dimension	512
Mini Batch Size	32
Dropout	0.5
Optimizer	SGD (lr=0.0009, momentum=0.9)
Criterion Function	Cross Entropy
Num. of Features	531

Table 4. Parameter Setting of Importance LSTM

6. Simulation results

In this part, we discuss the simulation results of our models.

There are four parts. Comparison between MLP and LSTM, followed by their performances on each class. Performance of importance MLP and importance LSTM, followed by their performance on each class.

6.1. MLP and LSTM

Comparison over accuracy and loss between MLP and LSTM without feature selection are shown in Figure 3 and Figure 4. From Figure 3, we can see the accuracy between MLP and LSTM are very good and stable, around 95% - 96%.

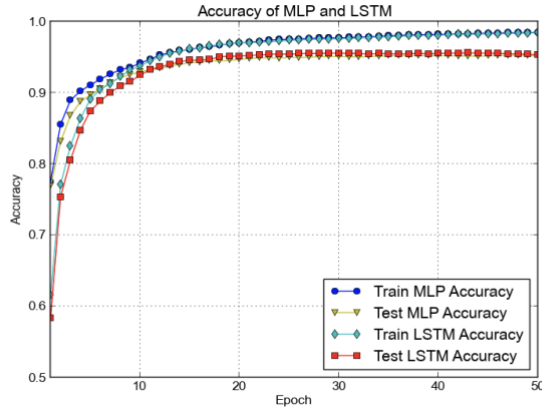


Figure 3. Accuracy of MLP and LSTM

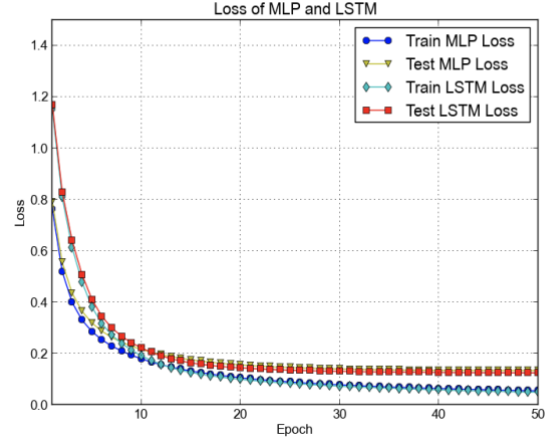


Figure 4. Loss of MLP and LSTM

6.2. MLP and LSTM on each class

From Figure 5 and Figure 6, LAYING and WALKING accuracy are very good in MLP and LSTM, but SITTING and WALKING DOWNSTAIRS accuracy are not good in MLP and LSTM. Both methods have troubles on SITTING prediction. SITTING prediction has a very large fluctuation while other activities gain better predictions.

In Figure 7, we know LSTM has better accuracy in SITTING and WALKING than MLP. MLP has better accuracy in STANDING and WALKING. Both of them have 100% accuracy in LAYING.

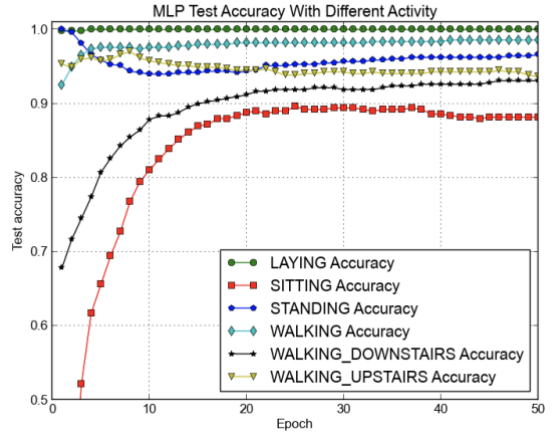


Figure 5. Accuracy of MLP on Each Class

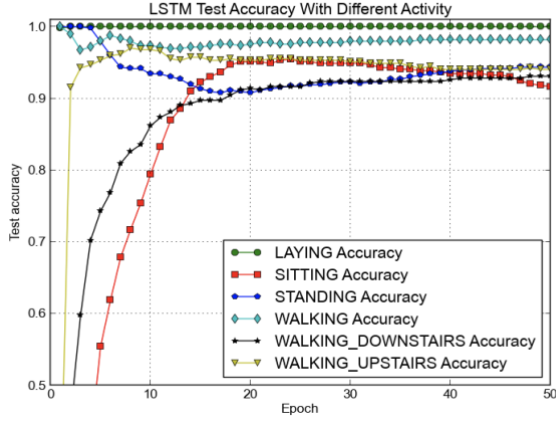


Figure 6. Accuracy with LSTM on Each Class

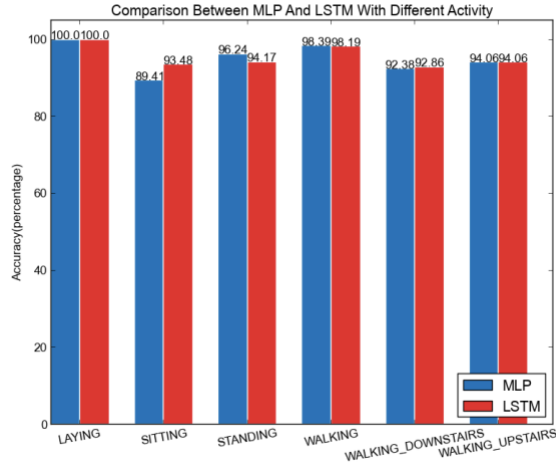


Figure 7. Comparison Between MLP and LSTM on Each Class

6.3. Feature selection of MLP and LSTM

From Figure 8, Importance MLP shows a good performance with 531 features rather than 561 features which will converge more faster and save memory with less features. From Figure 9, Importance LSTM has nearly same performance as LSTM.

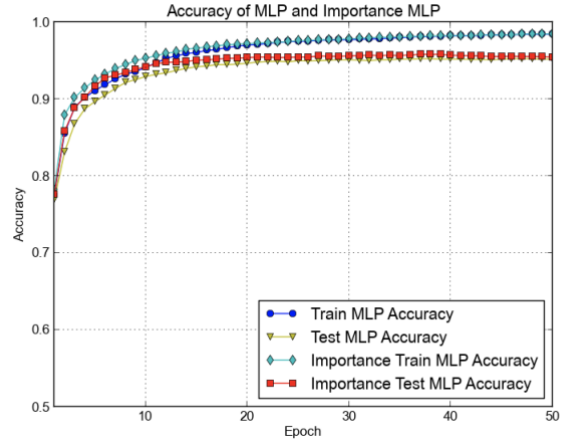


Figure 8. Comparison Between MLP and importance MLP

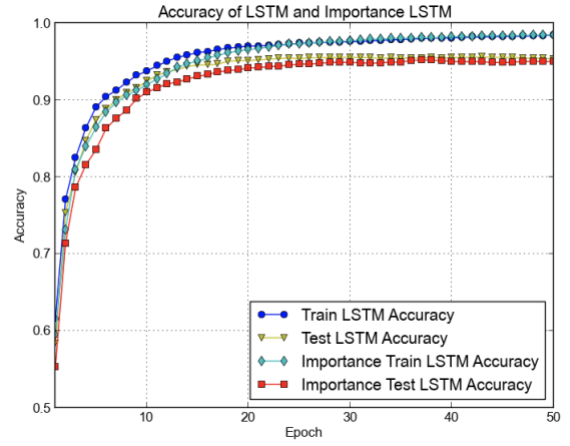


Figure 9. Comparison between LSTM and importance LSTM

6.4. Importance MLP and importance LSTM on each class

From the Figure 10 and Figure 11, we can see that, in MLP, test accuracy of LAYING, SITTING by importance feature selection are better than without feature selection, but STANDING, WALKING and WALKING UPSTAIRS are worse than without feature selection. In LSTM, test accuracy of STANDING and WALKING UPSTAIRS by importance feature selection are better than without feature selection, but SITTING, WALKING DOWNSTAIRS are worse than without feature selection.

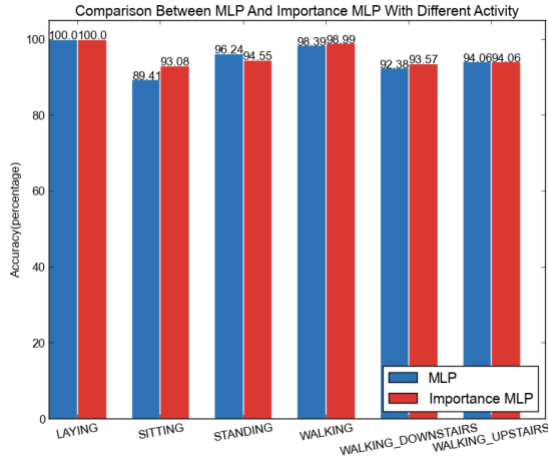


Figure 10: Comparison Between MLP and importance MLP on Each Class

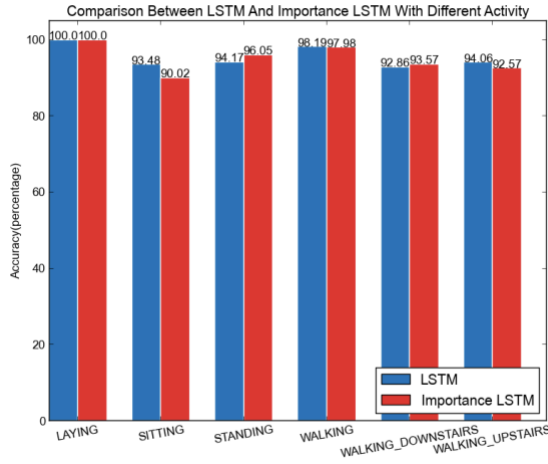


Figure 11: Comparison Between LSTM and importance LSTM on Each Class

7. Results analysis

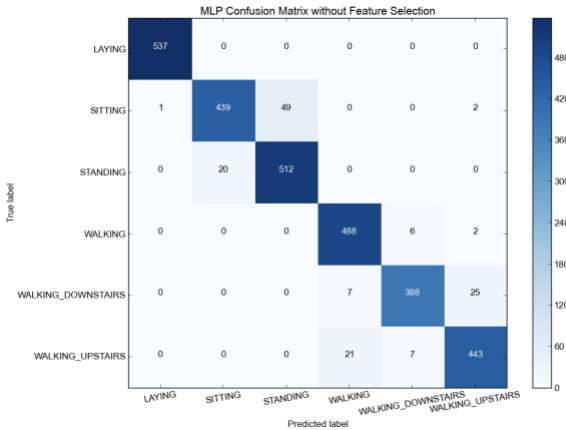


Figure 12: MLP Confusion Matrix Without Feature Selection

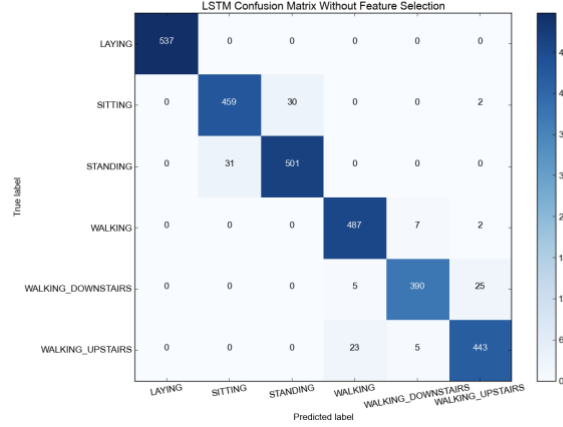


Figure 13: LSTM Confusion Matrix Without Feature Selection

Figure 12 and Figure 13 show confusion matrix from MLP and LSTM. From these confusion matrixes, we know there are confusion decision on SITTING and STANDING, also among WALKING, WALKING DOWNSTAIRS and WALKING UPSTAIRS. LAYING prediction is very good by MLP and LSTM. The tie-breaking accuracy is the SITTING accuracy and STANDING accuracy. If high accuracy achieved in SITTING and STANDING, overall accuracy is high. If not, overall accuracy will not good.

Besides importance feature selection, we also try Mutual Information Feature Selection which is a greedy selection of the features that takes both the mutual information with respect to the output class and with respect to the already-selected features into account. Finally, we get test accuracy around 93% -95%. We also try feature selection by F-Score, which considers both precision and recall of the test to compute score. The test accuracy of F-Score is around 91% - 94%.

Before implementing MIFS and F-Score feature selection, we think there is much information between features. By mutual information feature selection, we will get a good result. In fact, there are no improvement by this way. The main reason we think is that all feature values are already normalized, so they all in $[-1, 1]$. Many information loses due to normalization

Table 5 shows comparison test accuracy of our approaches and Nakano approaches [1]. CNN has better accuracy on WALKING, WALKING DOWNSTAIRS and WALKING UPSTAIRS. LSTM has better accuracy at SITTING. For STANDING, MLP has better accuracy. For LAYING, MLP, LSTM, importance MLP and importance LSTM have 100% accuracy. Overall, importance MLP has best accuracy 95.83%.

8. Conclusion and future work

Our two method MLP and LSTM has already show a

Activity	Other Approaches		Our Approaches			
	KNN	CNN	MLP	LSTM	Importance MLP	Importance LSTM
WALKING	0.9778	0.9980	0.9839	0.9819	0.9899	0.9798
WALKING UPSTAIRS	0.9002	0.9703	0.9406	0.9406	0.9406	0.9257
WALKING DOWNSTAIRS	0.7857	0.9905	0.9238	0.9286	0.9357	0.9357
SITTING	0.9304	0.7739	0.8941	0.9348	0.9308	0.9002
STANDING	0.9304	0.8909	0.9624	0.9417	0.9455	0.9605
LAYING	0.9944	0.8100	1.0000	1.0000	1.0000	1.0000
Overall	0.9012	0.9056	0.9525	0.9559	0.9583	0.9518

Table 5. Comparison between our approaches and other approaches

good performance with 561 features comparing with other methods such as CNN and KNN. They arrived 0.9525 and 0.9559 of test accuracy separately.

Considering there are too many features in our dataset (561) which converge slowly and large storage, we made improvements by selecting fewer features logically to save time and storage. We used importance in Random Forest Classifier to select features. We achieve 0.9583 accuracy by importance MLP.

In the future, we may try other models like SVM and feature selection like correlation feature selection to improve the accuracy.

References

- [1] K. Nakano and B. Chakraborty, "Effect of dynamic feature for human activity recognition using smartphone sensors" 2017 IEEE 8th International Conference on Awareness Science and Technology.
- [2] R. Hussein, J. Lin, K. Madden, and Z. J. Wang, "Robust Recognition of Human Activities using Smartphone Sensor.
- [3] H. Rahmani, A. Mian, and M. Shah, "Learning a deep model for human action recognition from novel viewpoints," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1-1, 2017.
- [4] D. Anguita, A. Ghio, L. Oneto, et al., "A public domain dataset for human activity recognition using smartphones". In: Proceedings of the European symposium on artificial neural networks (ESANN), Bruges, pp.427-442, 2013.