Hierarchical Classification Using ML/DL for Sussex-Huawei Locomotion-Transportation (SHL) Recognition Challenge

Yi-Ting Tseng
National Chiao Tung University
Hsinchu 300, Taiwan
ss920386.cs04@nctu.edu.tw

Hsien-Ting Lin
National Chiao Tung University
Hsinchu 300, Taiwan
scott41216.cs03@g2.nctu.edu.tw

Yi-Hao Lin
National Chiao Tung University
Hsinchu 300, Taiwan
hank1377.cs02g@nctu.edu.tw

Jyh-Cheng Chen*
National Chiao Tung University
Hsinchu 300, Taiwan
jcc@cs.nctu.edu.tw

ABSTRACT

In this paper, our team, SensingGO, presents a hierarchical classifier for Sussex-Huawei Locomotion-Transportation (SHL) recognition challenge. We first separate the original data into motorized activities and non-motorized activities in the first layer of the classifier by using accelerometer data. For the non-motorized activities, we calculate auto-correlation values with accelerometer data as input features. For the motorized activities, we take magnetometer and barometer with mean, maximum, standard deviation values as input features. Finally, we integrate the recognition results of each layer of the classifier, and the average F1-score is 50% to the validation data.

CCS CONCEPTS

• Computing methodologies → Supervised learning by classification.

KEYWORDS

datasets, neural networks, machine learning, activity recognition

ACM Reference Format:

Yi-Ting Tseng, Hsien-Ting Lin, Yi-Hao Lin, and Jyh-Cheng Chen. 2020. Hierarchical Classification Using ML/DL for Sussex-Huawei Locomotion-Transportation (SHL) Recognition Challenge. In *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers (UbiComp/ISWC '20 Adjunct), September 12–16, 2020, Virtual Event, Mexico.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3410530.3414347

*Corresponding author

UbiComp/ISWC '20 Adjunct, September 12–16, 2020, Virtual Event, Mexico

ACM ISBN 978-1-4503-8076-8/20/09…$15.00
https://doi.org/10.1145/3410530.3414347

1 INTRODUCTION

Human activity recognition (HAR) is an interesting application which aims at identifying the mean of activities a person is posing. There are many prior works [2, 4, 5, 7, 8] to address HAR challenge, including computer vision, wireless signal, wearable devices, and smartphones. The Sussex-Huawei Locomotion-Transportation (SHL) recognition challenge in 2020 provides Sussex-Huawei Locomotion-Transportation (SHL) dataset [6, 11], which contains seven types of sensor data from a smartphone to recognize eight modes of human activities, such as still, walk, run, bike, car, bus, train, and subway. The sensor data consist of the data collected from accelerometer, gravity sensor, gyroscope sensor, linear accelerometer, magnetic force, and barometer. The raw sensor data are sampled with 100 Hz and recorded from a person and four phone locations, such as bag, hips,torso, and hand.

In this paper, we present a hierarchical classifier (team name: SensingGO) to the SHL recognition challenge. The classifier consists of three modules: (1) motion classifier in the first layer, (2) non-motorized classifier in the second layer, and (3) motorized classifier in the second layer. By the feature analysis, we find that inertial sensor data are promising to separate the original data into motorized and non-motorized activities. Then, the variation of air pressure and magnetic field have a positive correlation with moving speed between motorized activities. Auto-correlation values of accelerometer data provides regularity features for the non-motorized activities. We apply XGBT and a fully-connected neural network as classifiers for motorized and non-motorized activities, respectively. Finally, our proposed method performs 50% F1-score in average with validation data.

In the following, we present our method with more details, and elaborate our accuracy.

2 DATASET

The overall data belongs to the SHL recognition challenge named as Sussex-Huawei Locomotion-Transportation (SHL) dataset [6, 11] including train data, validation data and test data. Both train data and validation data contain four phone locations (bag, hips, torso, hand). The train data is recorded by a single participant over 59 days, and the validation data is recorded by two participants over 6 days. The dataset contains the following sensor data: acceleration, gravity, rate of turn, linear acceleration, magnetic field, orientation and pressure. The dataset is labeled with the following eight activities: still, walking, run, bike, car, bus, train, and subway. Three datasets are segmented in non-overlapping windows with 5 seconds length and the order is guaranteed in 5 seconds. The ratio of each label is not the same. Statically, the label Still accounts for the highest ratio which is 20% in all labels. On the other hand, the label Run just accounts for 1.9%.

3 METHOD

3.1 Overview

We design a hierarchical model to classify SHL dataset to apply different features and models to retrieve the final results gradually. The labels we use for model training are decided by focusing on the mode of the whole 500 samples within a single record instead of considering all samples within 5 seconds. The reason lies in our observation that 500 successive labels sampled in 5 seconds of both train data and validation data seldom differ, which means participants rarely transfer to other activity in 5 seconds. There are only 416 records with labels including different activities from the total 784288 records.

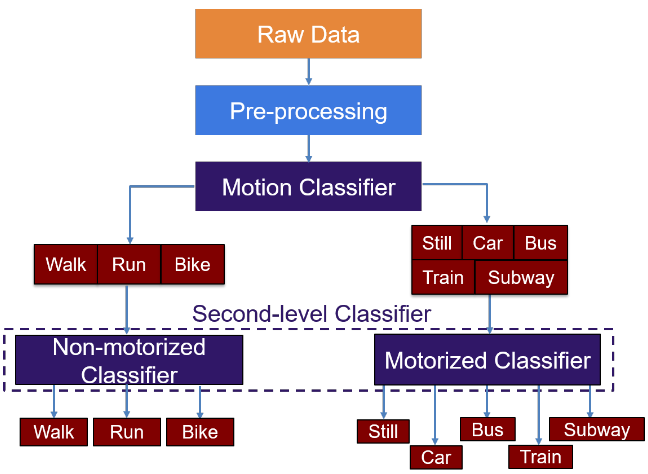

Figure 1: Overview of proposed hierarchical classifiers.

In this paper, we propose a 2-layer architecture consisting of three models. As shown in Fig. 1, all data pass through the first-level classifier, the motion classifier, to be roughly classified into two groups. The first group is supposed to be within walking, running and biking, while the second group shall be within still, car, bus, train and subway. There are two second-level classifiers, non-motorized classifier and motorized classifier, that further classify those two groups to give out the final results. The motion classifier and motorized classifier both use the XGBT model provided by Python API [1]. On the other hand, the non-motorized classifier uses a 3-layer MLP which has 80, 10 and 3 neurons for the first, second and third layer, respectively, and a dropout layer is appended before the third layer.

In the following sections, we will give more detailed introduction to explain the features and models we apply for those classifiers.

3.2 First-level Classifier - Motion Classifier

The motion classifier, which predicts whether the input data is within walking, running and biking or the rest, is built up with XGBT and is used as the first-level classifier. By observing the statistics of sensors data, we found that data may be roughly classified into two groups. A simple example is shown in Fig. 2. Fig. 2 shows the boxplot presenting the distribution of standard deviations of each acceleration record with 500 samples. The figure suggests that the class walking, run and bike have more different distribution from the rest, and this is also observed from our other experiments using different statistics of acceleration and linear acceleration. Therefore, the basic idea of the motion classifier is to separate the data of walking, running and bike from that of still, car, bus and subway. To reduce the impact of orientation change caused by human movements, we calculate force of acceleration and linear acceleration from the original components in three axes x, y and z. In order to extract potential information, we calculate statistics like standard deviation, maximum and minimum of acceleration and linear acceleration within five hundred samples for each record. Finally, we utilize those statistics to train the motion classifier with XGBT.

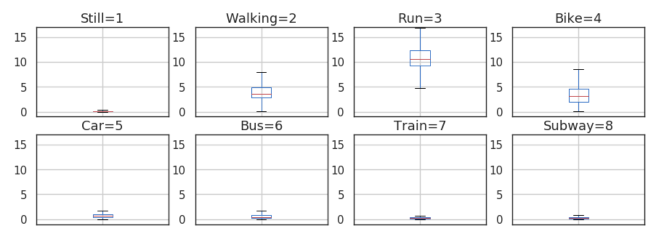

Figure 2: Boxplots of the standard deviation of acceleration for all classes.

3.3 Second-level Classifiers - Motorized and Non-motorized Classifier

After passing through the first level of classifier, the data is separated into two groups. The data in one group is predicted to be within non-motorized labels including walking, running and biking, and the data in the other group is predicted to be possible for still and other motorized labels including car, bus, train and subway. Still is grouped together with motorized labels other than the non-motorized ones since users' are usually still on motorized vehicles, and the observed behaviors of sensors from a still user are hence more similar to those of a user on a motorized vehicular.

Before further classifying these two groups of data, we would like to convert the coordinate system of data collected from linear accelerometer. As suggested by [12], different orientations of device will change the acceleration even if the user is doing the same thing in the same way, and that causes an issue in activity recognition. Therefore, the tri-axial linear accelerations denoted as LA_x, LA_y and LA_z are transformed to vertical and horizontal linear accelerations LA_v and LA_h. LA_v points toward the center of the Earth like gravity does, and LA_h is perpendicular to it. The transformation is calculated together with the gravity in three axes x, y and z, denoted as G_x, G_y and G_z. The angle θ between gravity and user's linear acceleration is calculated first, and then the projection can be applied to retrieve LA_v and LA_h. The procedure is shown in the following equations:

$$\theta = \arccos\left(\frac{G_x LA_x + G_y LA_y + G_z LA_z}{\sqrt{{G_x}^2 + {G_y}^2 + {G_z}^2}\sqrt{{LA_x}^2 + {LA_y}^2 + {LA_z}^2}}\right) \tag{1}$$

$$LA_v = \sqrt{{LA_x}^2 + {LA_y}^2 + {LA_z}^2} \cdot \cos\theta \tag{2}$$

$$LA_h = \sqrt{{LA_x}^2 + {LA_y}^2 + {LA_z}^2} \cdot \sin\theta \tag{3}$$

Next, to classify between walking, running and biking, we apply *auto-correlation coefficient function* [3] saying that given measurements $X_1, X_2, \ldots, X_N$ at time $t_1, t_2, \ldots, t_N$, the lag k auto-correlation function is defined as

$$r_k = \frac{\sum_{i=1}^{N-k}(X_i - \overline{X})(X_{i+k} - \overline{X})}{\sum_{i=1}^{N}(X_i - \overline{X})}. \tag{4}$$

Auto-correlation represents the similarity between the values of the same variables over successive time intervals and is useful for finding repeating patterns from data. We compute the auto-correlation coefficients of LA_v for each record with 500 samples. To reduce the input dimension for later training purpose, we reduce the number of observed lag k auto-correlation function from 500 to 250. That is, we only consider r_k where $0 \leq k \leq 250$.

The auto-correlation coefficients of LA_v of some records while walking, running and biking are visualized in Fig. 3. It is observable that the patterns of LA_v while walking are much more irregular than that of LA_v while running and biking. Furthermore, speaking of the oscillation between positive and negative correlation, the auto-correlation coefficients of biking have a higher frequency than that of the running. Therefore, we believe auto-correlation coefficients are useful information to classify between walking, running and biking. We built a neural network which is a 3-layer MLP for the non-motorized classifier and the evaluation result will be shown in Sec. 4.

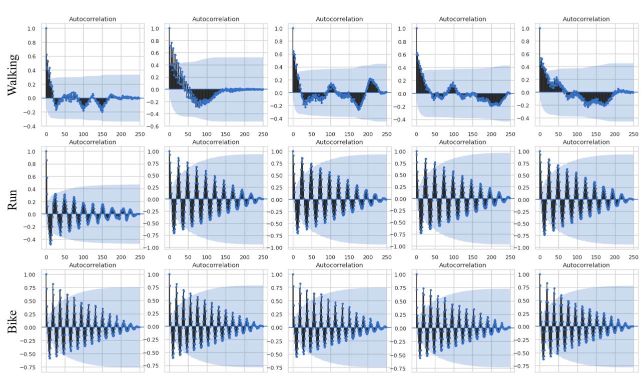

Figure 3: Auto-correlation of LA_v while walking, running and biking.

To classify the other group of data including still, car, bus, train and subway with the motorized classifier, since the users' motions are similar in these cases, we rely on other kinds of features instead of using acceleration as before. The first one is magnetic field. Vehicles like train and subway tend to have higher fluctuation in the magnetic field due to the material of carriage, the electricity power used for the vehicles, and other factors from the track [9]. As shown in Fig. 4 which is the boxplots of the variance of magnetic field calculated from the data labeled with still, car, bus, train and subway, the distributions of train and subway are obviously different from the others'.

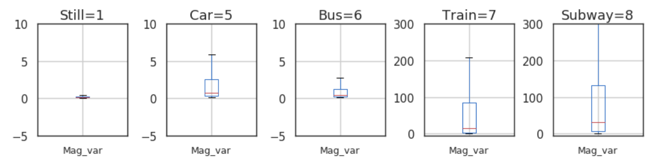

Figure 4: Boxplots of the variance of magnetic field with labels of still, car, bus, train and subway.

The second feature we pick is the air pressure measured by a barometer. The property we mainly utilize is that the volume of the space where the user is may affect the stability of measured air pressure [9]. In general, the larger the carriage, the stabler the air pressure. Fig. 5 and Fig. 6 show the variance and mean value of the pressure of the data labeled with still, car, bus, train and subway. The air pressure in a car has a larger variance than that in a bus or train since it has a smaller car body, which leads to larger fluctuation in air pressure. A still user outside of any vehicular measures the stablest air pressure. Moreover, since subway is underground, it has an average air pressure larger than other vehicles.

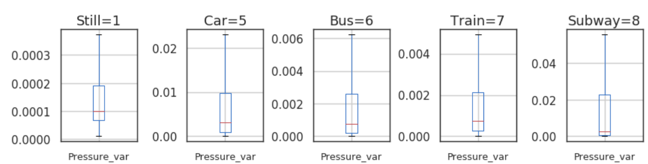

Figure 5: Boxplots of the variance of pressure with labels of still, car, bus, train and subway.

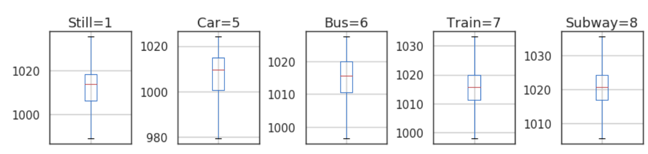

Figure 6: Boxplots of the average of pressure with labels of still, car, bus, train and subway.

Due to all the reasons mentioned above, we decide to let the motorized classifier that recognizes still, car, bus, train and subway rely on the statistics of measured values of magnetic field and air pressure. The model is built by XGBT and the evaluation results will be given in Sec. 4.

4 EVALUATION

All of our experiments are conducted on a PC with a 16-core CPU, 16GB RAM, and an Nvidia GTX 1080 GPU. The overall training time takes about 46 minutes. To be more detailed, the motion classifier and the motorized classifier both utilize XGBT and their training time are 18 seconds and 83.5 seconds, averaged by 5-time experimental results respectively. It can be inferred that the overall training time is mainly contributed by the non-motorized classifier which is a 3-layer MLP. To test its training time, we trained the MLP for 30 epochs for 5 times, and found it averagely early-stopped (when it fails to progress for more than 5 epochs) at epoch 14. Since each epoch takes about 190 seconds, the average training time for the non-motorized classifier can be said to be about 2660 seconds. We evaluate our models with the validation data provided by the SHL recognition challenge. As for the testing time, we ran the models with validation data for 5 times, and the average testing time for the motion classifier, the motorized classifier and the non-motorized classifier are 0.1 second, 0.1 second, and 3.4 seconds, respectively. Since the second-level classifiers can be run in parallel, the averaged overall testing time is about 3.5 seconds.

For the motion classifier, it can successfully separates non-motorized activities from motorized activities with an F1-score of 88.6%. As for the second-level classifiers, the non-motorized classifier using only coefficients of auto-correlation function as features can classify within the non-motorized classes including walking, running and biking with an F1-score of 84.7%. The good performance implies that the coefficients of auto-correlation function may be powerful features to recognize non-motorized activities. On the other hands, the motorized classifier achieves only 60.3% F1-score. We further explore the classification result of the motorized classifier, and the misclassifications mainly happen between car and bus and between train and subway. While using magnetic field feature singly can separate car/bus from train/subway with an F1-score of 83.2%, the unideal overall performance may imply that air pressure measured by barometers may not be a strong enough feature to distinguish between car and bus and between train and subway. The exploration of more useful features and the improvement of the motorized classifier may be our future work.

The confusion matrices of the classification results output from the first-level classifier and second-level classifiers are shown in Fig. 7 and Fig. 8, respectively. Fig. 7 shows that the motion classifier can successfully recognize 65% of non-motorized activities and that only 1% of the motorized activities are wrongly predicted to be non-motorized. As for the motorized activities, though the 99% of them can be predicted correctly, 35% of the predicted motorized activities are misclassifications. This means that our proposed motion classifier is unable to recognize these 35% of non-motorized activities.

We can further explore the classification results of the first-level classifier. Fig. 8 shows that the motion classifier confuses with data labeled with bike, and the data labeled with bus and train.The reason would be that the patterns of inertial sensor values were irregular for these vehicles, because users on these vehicles would be more likely to be affected and shaken by the movements of the vehicles.

As for the second-level classifier, Fig. 8 shows that the non-motorized classifier can not identify the data labeled with run, and the data labeled with bike, because their behaviors are similar.

Data within motorized activities that the motorized classifier identifies retains about 50% accuracy. The first reason is that the data that are labeled with still are more possibly to be misclassified with those labeled with bus and car. This would result from that users in car and bus may frequently be still when the vehicles stop due to the traffic lights or bus stations. The second reason would be that the mobility of car, bus, train and subway are similar which makes the variations of sensor data close and similar.

More detailed performance metrics of the three classifiers and the overall performance are presented in Table. 1.

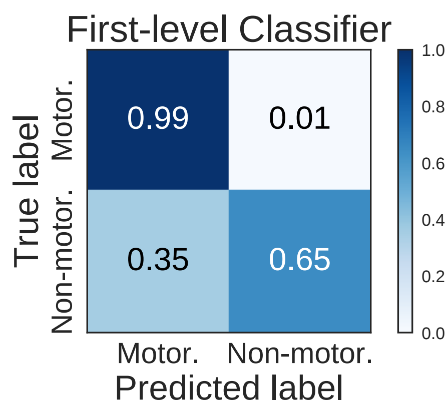

Figure 7: Confusion matrix of first-level classifier evaluation result.

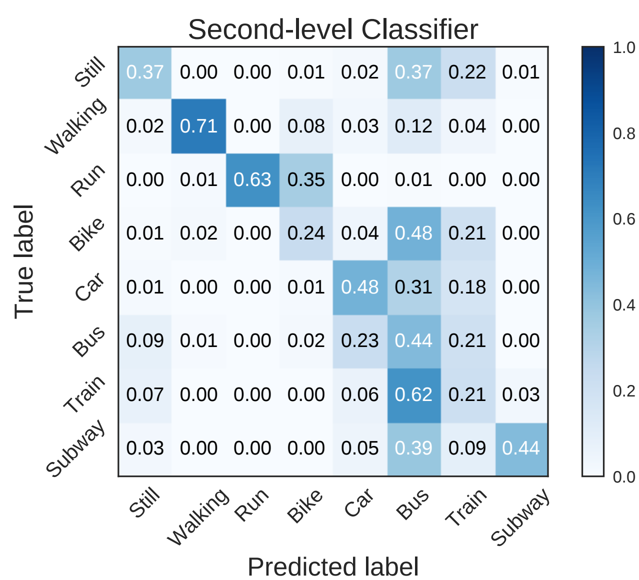

Figure 8: Confusion matrix of second-level classifier evaluation result.

Table 1: Performance Metrics.

	Precision	Recall	F1-score
Motion classifier	90.1%	89.3%	88.6%
Non-motorized classifier	87.8%	84.3%	84.7%
Motorized classifier	60.2%	60.7%	60.3%
Overall	64.7%	43.4%	49.9%

5 CONCLUSION

In this paper, we propose a hierarchical classifier with three models using different architectures and features to the SHL recognition challenge. Based on our observation of different sensors' behavior, we firstly propose the motion classifier to separate all activities into non-motorized ones and motorized ones. Then, the non-motorized classifier makes use of auto-correlation values of linear accelerometer data to distinguish each label from non-motorized activities. Finally, we treat pressure and magnetic field as inertial features in the motorized classifier to recognize still and other different motorized activities. After a series of experiments, our method achieves about 50% F1-score in average with validation data.

The recognition result for the testing dataset will be presented in the summary paper of the challenge [10].

ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Science and Technology of Taiwan under grant numbers MOST 109-2218-E-009-004, MOST 108-2221-E-009-042-MY3, and MOST 108-2218-E-009-028.

REFERENCES

[1] Released: Jun 7, 2020. *Python Package Index PYPI: xgboost*. https://pypi.org/project/xgboost/

[2] Jacopo Biancat, Chiara Brighenti, and Attilio Brighenti. 2014. Review of Transportation Mode Detection techniques. *EAI Endorsed Transactions on Ambient Systems* 1, 4 (2014), e7.

[3] George Edward Pelham Box and Gwilym Jenkins. 1990. *Time Series Analysis, Forecasting and Control*. Holden-Day, Inc., USA.

[4] Chhavi Dhiman and Dinesh Kumar Vishwakarma. 2019. A review of state-of-the-art techniques for abnormal human activity recognition. *Engineering Applications of Artificial Intelligence* 77 (2019), 21–45.

[5] Shih-Hau Fang, Hao-Hsiang Liao, Yu-Xiang Fei, Kai-Hsiang Chen, Jen-Wei Huang, Yu-Ding Lu, and Yu Tsao. 2016. Transportation modes classification using sensors on smartphones. *Sensors* 16, 8 (2016).

[6] Hristijan Gjoreski, Mathias Ciliberto, Lin Wang, Francisco Javier Ordonez Morales, Sami Mekki, Stefan Valentin, and Daniel Roggen. 2018. The university of sussex-huawei locomotion and transportation dataset for multimodal analytics with mobile devices. *IEEE Access* 6 (2018), 42592–42604.

[7] Adrian C Prelipcean, Gyözö Gidófalvi, and Yusak O Susilo. 2017. Transportation mode detection–an in-depth review of applicability and reliability. *Transport Reviews* 37, 4 (2017), 442–464.

[8] Xing Su, Hernan Caceres, Hanghang Tong, and Qing He. 2016. Online travel mode identification using smartphones with battery saving considerations. *IEEE Transactions on Intelligent Transportation Systems* 17, 10 (2016), 2921–2934.

[9] Xing Su, Hernan Caceres, Hanghang Tong, and Qing He. 2016. Online Travel Mode Identification Using Smartphones With Battery Saving Considerations. *IEEE Transactions on Intelligent Transportation Systems* (03 2016), 1–14. https://doi.org/10.1109/TITS.2016.2530999

[10] L. Wang, H. Gjoreski, M. Ciliberto, P. Lago, K. Murao, T. Okita, and D. Roggen. 2020. Summary of the Sussex-Huawei locomotion-transportation recognition challenge 2020. *Proceedings of the 2020 ACM International Joint Conference and 2020 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (2020).

[11] Lin Wang, Hristijan Gjoreski, Mathias Ciliberto, Sami Mekki, Stefan Valentin, and Daniel Roggen. 2019. Enabling reproducible research in sensor-based transportation mode recognition with the Sussex-Huawei dataset. *IEEE Access* 7 (2019), 10870–10891.

[12] Rong Yang and Baowei Wang. 2016. PACP: A Position-Independent Activity Recognition Method Using Smartphone Sensors. *Information* 7 (12 2016), 72. https://doi.org/10.3390/info7040072