A Windowless Approach to Recognize Various Modes of Locomotion and Transportation

Promit Basak
Electrical and Electronic Engineering
University of Dhaka
Dhaka, Bangladesh
basakpromit@gmail.com

Shahamat Mustavi Tasin
Electrical and Electronic Engineering
University of Dhaka
Dhaka, Bangladesh
tasin.mustavi@gmail.com

A.H.M. Nazmus Sakib
Electrical and Electronic Engineering
University of Dhaka
Dhaka, Bangladesh
nazmussakib2970@gmail.com

Syed Doha Uddin
Electrical and Electronic Engineering
University of Dhaka
Dhaka, Bangladesh
doha.yeamin@gmail.com

Md Atiqur Rahman Ahad
Electrical and Electronic Engineering
University of Dhaka, Bangladesh
Department of Media Intelligent
Osaka University, Japan
Dhaka, Bangladesh
atiqahad@du.ac.bd

ABSTRACT

Detecting modes of transportation through human activity recognition is important in the effective and smooth operation of smartphone applications or similar portable devices. However, the effectiveness of such tasks depends on the nature and type of data provided, and it can often become quite challenging. “SHL recognition challenge 2021” is an activity recognition challenge that aims to detect eight modes of locomotion and transportation. The dataset in this challenge was based on radio data, including GPS reception, GPS location, Wi-Fi reception, and GSM cell tower scans. The objective was to create a model that was able to recognize the modes in a user-independent manner. In this paper, our team (Team Nirban) gives an appropriate summarization of our methods and approach to the challenge. We processed the data, extracted various features from the dataset, and selected the best ones, which helped our model to be generative and user-independent. We exploited a classical machine learning based approach and achieved 93.4% accuracy and 89.6% F1 score on the training set using 10-fold cross-validation, as well as 62.3% accuracy on the provided validation set.

CCS CONCEPTS

• **Human-centered computing** → **Smartphones**; • **Computing methodologies** → **Supervised learning by classification**; **Classification and regression trees**.

KEYWORDS

Activity recognition; Locomotion; Sensor data; Statistical features; Random Forest

ACM Reference Format:
Promit Basak, Shahamat Mustavi Tasin, A.H.M. Nazmus Sakib, Syed Doha Uddin, and Md Atiqur Rahman Ahad. 2021. A Windowless Approach to Recognize Various Modes of Locomotion and Transportation. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers (UbiComp-ISWC '21 Adjunct), September 21–26, 2021, Virtual, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3460418.3479385

1 INTRODUCTION

In this era of modern technologies, the most common but one of the most useful devices is the smartphone. Numerous researches can be done using the data from various sensors inside a smartphone, one of which is human activity recognition. Human activity recognition studies are mostly focused on walking, running, swimming, and other daily activities. With increasing technology and science, and computational power in the world, multiple HAR works are done using the data from different sensors [1]. Using those data, it is possible to recognize transportation activities which is a section of HAR. Transportation activity recognition has a great influence on multiple factors such as parking spot detection, suggestion on traffic routes, public transport congestion prediction, analysis of road conditions, faster delivery route and time suggestions, and so on.

The SHL recognition challenge 2021 dataset contains radio data of eight modes of locomotion [3], [17]. In previous challenges, datasets contained inertial sensor data of smartphones which were based on motion sensors. Wearable device sensor data is dependent on specific locations of that device that has been worked on [20], [11]. This year with radio data, the main challenge was to find the accurate location. Radio data gives an approximate position of a place at best, but it can be off by 9-10 meters on average (up to 90-100 meters in some cases) [7]. To overcome this problem, we have emphasized on locations that had higher accuracy. The dataset

UbiComp-ISWC '21 Adjunct, September 21–26, 2021, Virtual, USA

ACM ISBN 978-1-4503-8461-2/21/09…$15.00
https://doi.org/10.1145/3460418.3479385

also contained cells, GPS, Wi-fi data from where we extracted key features to train our model.

Our team approached the challenge by working on the raw data first, and preprocessing it properly to use it in models. We extracted key features from the data to have better accuracy of the model. Not all features necessarily helped and some tend to overfit the model, so to generalize it, we chose the appropriate features. We used a random forest classifier because of its computational efficiency to train the data. Figure 1 describes the basic workflow in brief.

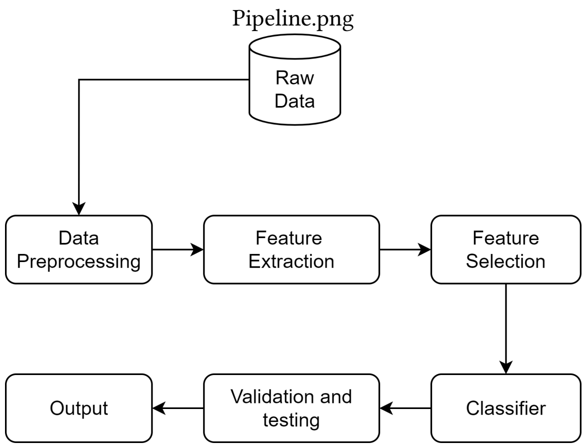

Figure 1: A Basic structure of our proposed method

The contents of the paper are organized as follows: In the 2nd section, a broad and helpful description of the dataset is provided. Proper exploration of the dataset and analysis is also a part of this section. In section 3, we explain our proposed method for this challenge. The necessary preprocessing of data, feature extraction and selection, model classification are explained here. The results and evaluations of our method are stated in section 4. The limitations and further improvement in our work are discussed in the discussion which is section 5. Finally, the limitations and conclusions that are drawn in section 6.

2 CHALLENGE DATASET

The dataset for the SHL recognition challenge 2021 is used in this paper, which consists of eight modes of locomotion and transportation activities in a user-independent manner based on radio data, including GPS reception, GPS location, Wi-fi reception, and GSM cell tower scans. There are eight activities in total, and they are:

(1) Still
(2) Walking
(3) Run
(4) Bike
(5) Car
(6) Bus
(7) Train
(8) Subway

From Figure 2, we can see that the data is much imbalanced for the challenge. Specifically, activity label 3 (Run) has so much fewer data to train on. This is one of the biggest problems to overcome.

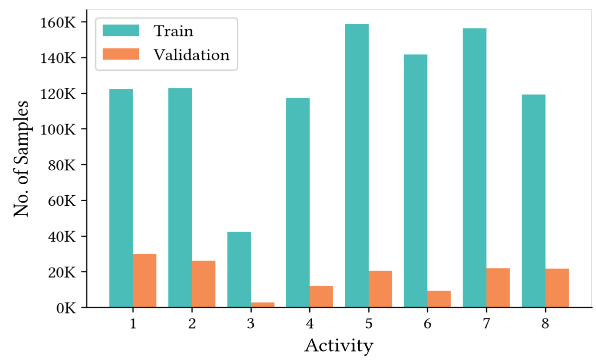

Figure 2: Sample distribution among different classes

The train, the validation, and the test parts each contain five text files. The files and their contents in the training data are:

(1) Location.txt
  - Epoch time [ms]
  - Accuracy of the location [m]
  - Latitude [degrees]
  - Longitude [degrees]
  - Altitude [m]
(2) GPS.txt
  - Epoch time [ms]
  - Number of available satellites
  - SNR of available satellites
  - Azimuth of available satellites
  - Elevation of available satellites
(3) WiFi.txt
  - Epoch time [ms]
  - Variable number of Wi-Fi data including BSSID, SSID, RSSI, Frequency [MHz] and capabilities
(4) Cells.txt
  - Epoch time [ms]
  - Variable number of data points based on available GSM, LTE, or WCDMA signals
(5) Label.txt
  - Epoch time [ms]
  - Labels

It is to be noted that all the sensors are asynchronously sampled. The sampling rate is roughly 1 Hz but is time-varying for each sensor. Note that, depending on the condition of the satellite and cell, one sensor may receive no signal at all at a certain interval, and thus no data recorded.

3 METHOD

We opted for classical machine learning methods for the "SHL recognition challenge 2021". The sliding window approach is one of the most popular approaches used with time series data as given in this challenge [8]. But in cases, where the frequency is too low, or the system is infrequently switching between states, it may not be possible to increase the window size as needed to identify complex activities [2]. In this dataset, the average sampling rate was close to 1Hz, which is very low. So, we opted for a windowless approach instead of the traditional windowed approach. Figure 3 depicts the summary of our methodology.

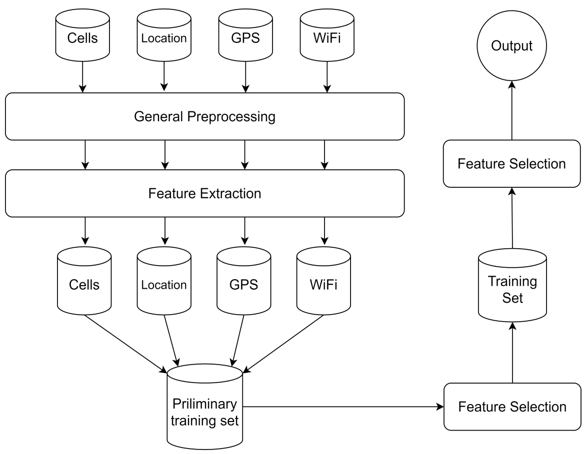

Figure 3: Our proposed methodology

3.1 Pre-processing

The given training data contained five different files: Cells.txt, GPS.txt, Label.txt, Location.txt, and WiFi.txt. Out of the five files, Label.txt contained the labels for all the other four data files. The data files were inconsistent, and each row could contain an indefinite number of columns. That's why we used text mining techniques to parse all the data in a single structured format without losing any portion of the data. One of the challenges of the dataset was different timestamps in different files, which created a problem while merging the data from different files. So, in this step, we processed the fixed timestamps of the Label.txt so that they represent a continuous period of time that solved the issue mostly. We also used a median filter for removing outliers, dropped duplicated and non-numerical rows from the data, and arranged them properly to make them suitable for the next step, which is feature extraction.

3.2 Feature Extraction

Unlike the previous installments of this challenge [18], [15], [16] this year, the data is based on radio sensors, including GPS reception, GPS location, Wi-Fi reception, and GSM cell tower scans. So, the increased number of data variety makes it more important to extract related features that are highly correlated to the target labels of this challenge. We extracted features individually from each of the data files and then merged them all. Hence, we will be describing them here one by one:

Cells: Cells.txt file contains cell tower scan data for three cellular communication standards, LTE, WCDMA, and GSM. For each standard of communication, we extracted some statistical features (mean, median, standard deviation, and sum) for different cellular properties, i.e., signal level, signal strength in ASU and dBm, number of registered connections. A total of 24 features were extracted from cell reception data.

Location: GPS location data, including latitude, longitude, altitude, and location accuracy, were there in this file. We used the Haversine formula, well-known for determining the distance between two points on the globe [10]. The Haversine formula is given below:

$$d = 2r \arcsin \sqrt{\sin^2(\frac{\phi_2 - \phi_1}{2}) + \cos\phi_1 \cos\phi_2 \sin^2(\frac{\psi_2 - \psi_1}{2})} \quad (1)$$

where, d is the distance between two coordinates of latitudes and longitude of (ϕ, ψ) and r is the radius of the earth.

From the change of distance, we calculated velocity, acceleration, and jerk by differentiating the distance with respect to time repeatedly. These features played the most important role in determining the target activity.

$$v = \frac{\Delta s}{\Delta t} \quad (2)$$

$$a = \frac{\Delta v}{\Delta t} \quad (3)$$

$$j = \frac{\Delta a}{\Delta t} \quad (4)$$

where, s, v, a and j corresponds to displacement, velocity, acceleration and jerk accordingly.

GPS: We extracted several features from the GPS reception data; most of them are different flavors of Dilution of Precision (DOP), i.e., PDOP, HDOP, VDOP, GDOP and TDOP [19]. DOPs are used to calculate how accurate GPS signal is. The lower value of DOPs points to the better accuracy of GPS. We can calculate DOPs from the following formulae:

$$GDOP = \sqrt{s_x^2 + s_y^2 + s_z^2 + s_t^2} \quad (5)$$

$$TDOP = \sqrt{s_t^2} \quad (6)$$

$$PDOP = \sqrt{s_x^2 + s_y^2 + s_z^2} \quad (7)$$

$$HDOP = \sqrt{s_x^2 + s_y^2} \quad (8)$$

$$VDOP = \sqrt{s_z^2} \quad (9)$$

Wi-Fi: WiFi.txt contained BSSID, SSID, RSSI, Frequency [MHz] and capabilities data. We extracted different statistical features for the number of Wi-Fi signals, strength, frequency, and security systems used in the Wi-Fi systems.

3.3 Feature Selection

Not all features are related to the goal of the challenge. Our model might learn unnecessary details from redundant and irrelevant features that do not apply to generalized cases. This may lead to overfitting our model too quickly. That's why we made use of different feature selection techniques that helped us get rid of features with negative impact. The used techniques include:

1. Chi-Square Test [13]: It is a numerical test that is used to measure the deviation from the expected distribution and is used to estimate whether the target label is independent of a feature.

2. Pearson Correlation Coefficient (PCC) [6]: PCC measures how strongly two variables are related to each other. It helps to remove redundant features that are linearly dependent.

3. Decision Tree Embedded Selection [12]: DT-based models use impurity measurements that can be utilized to select the most

Table 1: Comparison of performance among the four models

Model	Accuracy (%)	F1 score (%)	Validation Accuracy (%)
KNN	85.2	80.1	42.2
RF	93.4	89.6	62.3
ET	93.3	89.2	47.5
XGB	92	87.5	53.8

important features. It was more convenient to use for us as we were using DT-based models for our classification.

We used all three techniques and used the mean importance to select the most important 52 features for our proposed method.

4 RESULTS

The challenge dataset was very much imbalanced. Generally, Decision Tree-based models work well on determining human activity from an imbalanced dataset [5]. We have done 10-fold cross-validation on four different models, namely, K-Nearest Neighbor classifier (KNN), Random Forest classifier (RF), Extra Trees classifier (ET), and XGBoost classifier (XGB), and used the best model to predict on the validation set.

We explored some linear and statistical models such as the Logistic regression, and Naive Bayes classifier to set the baseline performance for the task. After that, we used three decision-tree-based models and one nearest neighbor model to evaluate which model works best on our method. We used 10-fold cross-validation and found out the Random Forest model gives the best result in the dataset with an accuracy of 93.4%. On the validation set, it obtains an accuracy of 62.3%.

As stated in Table 1, the Decision Tree based models did better than KNN model. Results obtained from ET and RF models are very close and RF performed slightly better than ET. Hence, we selected Random forest model to predict on test data.

Table 2 shows the activity wise performance of the RF model in the validation data. The table has some detailed information about precision, recall, F1 score, and support for each activity. And the overall accuracy for all the eight activities is recorded in the accuracy column.

In Figure 4, the confusion matrix of prediction on the validation data is shown. It is seen that activities 7 (Train) and 8 (Subway) have been very frequently confused with other activities. This shows that our model has struggled to identify these activities. Also, activity 3 (Run) was hardly identified because of a very fewer number of samples.

5 DISCUSSION

In the validation data, activities 3, 4, and 8 contained a much lower number of samples than that of other activities. The confusion matrix shows that these activities were frequently mislabeled or

Table 2: Activity-wise classification report for RF model on validation data (in %)

Activity ID	Precision	Recall	F1 score	Support	Accuracy
1	99	38	55	16996	62
2	87	84	85	31906	
3	15	6	8	3676	
4	27	99	42	4534	
5	64	83	72	39103	
6	41	50	45	13457	
7	47	10	16	12832	
8	56	27	36	6731	
Macro avg.	55	50	45		
Weighted avg.	67	62	60		

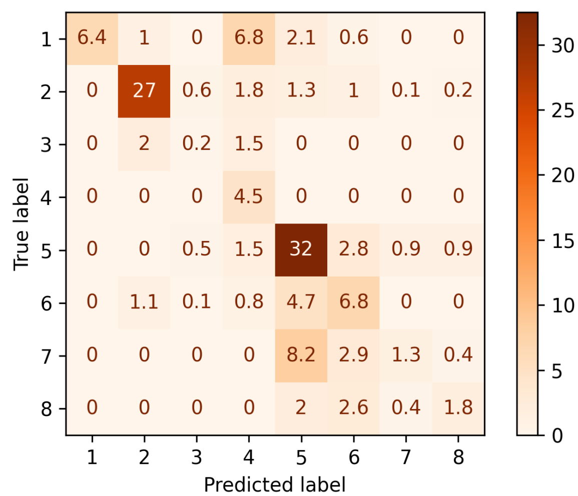

Figure 4: Confusion matrix obtained using RF model on the validation set

confused with other activities. So, the imbalance of the dataset had an impact on the result. Random under-sampling for major classes along with artificial oversampling for minor classes may be a solution for this problem. Also, there is a noteworthy difference between the median accuracy obtained from 10-fold cross-validation on the training set and the accuracy on the validation set. This shows that there was overfitting problem on the data during training. One of the reasons for overfitting was the inconsistency of data in the four data files. For any given timestamp, it was seen in many cases that the data is not present in all the files. So, a significant amount of data could not be utilized in this process. This also led to a repetition of the same data due to interpolation which is one of the root causes of overfitting. Another reason for overfitting might be that the training data was collected by user 1 while the validation and test data were collected by user 2 and user 3.

One of the drawbacks of the windowless system is that it considers a single timestamp to predict the activity. This drawback may lead to some wrong predictions for any instantaneous change in features. We can easily overcome this issue by using a median filter to filter out the odd predictions.

6 CONCLUSION

In our work, we have taken radio data collected with smartphones and extracted statistical features to train on different models. Among them, the Random Forest classifier performed the best based on 10-fold cross-validation result. Some important observations have been noted which can be addressed to get better outcomes. A deeper dive into cellular network data and Wi-Fi data may be fruitful. We hope to work on these issues and get better results in the future. Furthermore, we want to implement other models such as LightGBM [4], CatBoost [9] and end-to-end deep learning approaches on this dataset. The recognition result for the testing dataset will be presented in the summary paper of the challenge [14].

Table 3: Appendix

Used data	GPS reception, GPS location, Wi-Fi reception, and GSM cell tower scans.
Selected Features	Total number of detected and registered cellular signals for each of GSM, WCDMA and LTE, sum, mean and median of signal strength, ASU level, signal level.
	Location accuracy, distance, altitude, velocity, acceleration and jerk.
	PDOP, TDOP, GDOP< HDOP, VDOP and number of satellites.
	Mean, standard deviation and median of Wi-Fi strength, frequency, Signal-to-noise ration (SNR) and total number of Wi-Fi signals using TKIP and CCMP security protocol.
Programming language	Python 3
Libraries used	NumPy, Pandas, SciPy, Scikit-learn, xgboost, matplotlib, seaborn
Training and testing time	Training set (approx. 1.3 million data units): Approximately, 150 minutes for Data processing, Feature Extraction, Training, and Prediction.
	Test set (approx. 1.1 million data units): Approximately, 90 minutes for Data processing, Feature Extraction, and Prediction.
Machine specification (RAM, CPU)	RAM 12GB, Disk 107 GB, CPU: Intel Xeon @2.20 GHz, Cores: 2.

REFERENCES

[1] Md Atiqur Rahman Ahad, Anindya Das Antar, and Masud Ahmed. 2021. *IoT Sensor-Based Activity Recognition.* Vol. 173. Springer International Publishing, United States. https://doi.org/10.1007/978-3-030-51379-5

[2] Philippe Giguere and Gregory Dudek. 2008. Clustering Sensor Data for Terrain Identification using a Windowless Algorithm. In *Robotics: Science and Systems IV.* The MIT Press, United States, 25–32. https://doi.org/10.15607/RSS.2008.IV.004

[3] 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. https://doi.org/10.1109/ACCESS.2018.2858933

[4] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. LightGBM: a highly efficient gradient boosting decision tree. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems* 30 (12 2017), 3149–3157.

[5] Bartosz Krawczyk, Michał Woźniak, and Gerald Schaefer. 2014. Cost-sensitive decision tree ensembles for effective imbalanced classification. *Applied Soft Computing* 14 (1 2014), 554–562. https://doi.org/10.1016/j.asoc.2013.08.014

[6] Yaqing Liu, Yong Mu, Keyu Chen, Yiming Li, and Jinghuan Guo. 2020. Daily Activity Feature Selection in Smart Homes Based on Pearson Correlation Coefficient. *Neural Processing Letters* 51 (4 2020), 1771–1787. Issue 2. https://doi.org/10.1007/s11063-019-10185-8

[7] Krista Merry and Pete Bettinger. 2019. Smartphone GPS accuracy study in an urban environment. *PLOS ONE* 14, 7 (7 2019), 1–19. Issue 7. https://doi.org/10.1371/journal.pone.0219890

[8] Javier Ortiz Laguna, Angel García Olaya, and Daniel Borrajo. 2011. A Dynamic Sliding Window Approach for Activity Recognition. In *User Modeling, Adaption and Personalization,* Joseph A. Konstan, Ricardo Conejo, José L. Marzo, and Nuria Oliver (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 219–230.

[9] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. 2017. CatBoost: unbiased boosting with categorical features. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems* (Montréal, Canada) *(NIPS'18).* Curran Associates Inc., Red Hook, NY, USA, 6639–6649. http://arxiv.org/abs/1706.09516

[10] Nitin R.Chopde and Mangesh K. Nichat. 2013. Landmark Based Shortest Path Detection by Using A* and Haversine Formula. *International Journal of Innovative Research in Computer and Communication Engineering* 1 (4 2013), 298–302. Issue 2.

[11] Md Sadman Siraj, Md Ahasan Atick Faisal, Omar Shahid, Farhan Fuad Abir, Tahera Hossain, Sozo Inoue, and Md Atiqur Rahman Ahad. 2020. UPIC: user and position independent classical approach for locomotion and transportation modes recognition. 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* (New York, NY, USA). ACM, New York, NY, USA, 340—-345. https://doi.org/10.1145/3410530.3414343

[12] V. Sugumaran, V. Muralidharan, and K.I. Ramachandran. 2007. Feature selection using Decision Tree and classification through Proximal Support Vector Machine for fault diagnostics of roller bearing. *Mechanical Systems and Signal Processing* 21, 2 (2 2007), 930–942. Issue 2. https://doi.org/10.1016/j.ymssp.2006.05.004

[13] Ikram Sumaiya Thaseen and Cherukuri Aswani Kumar. 2017. Intrusion detection model using fusion of chi-square feature selection and multi class SVM. *Journal of King Saud University - Computer and Information Sciences* 29, 4 (10 2017), 462–472. Issue 4. https://doi.org/10.1016/j.jksuci.2015.12.004

[14] Lin Wang, Mathias Ciliberto, Hristijan Gjoreski, Paula Lago, Kazuya Murao, Tsuyoshi Okita, and Daniel Roggen. 2021. Locomotion and transportation Mode Recognition from GPS and radio signals: Summary of SHL Challenge 2021. *Proceedings of the 2021 ACM International Joint Conference and 2021 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* 11 (2021).

[15] Lin Wang, Hristijan Gjoreski, Mathias Ciliberto, Paula Lago, Kazuya Murao, Tsuyoshi Okita, and Daniel Roggen. 2019. Summary of the Sussex-Huawei locomotion-transportation recognition challenge 2019. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers* (London, United Kingdom) *(UbiComp/ISWC '19 Adjunct)*. ACM, New York, NY, USA, 849–856. https://doi.org/10.1145/3341162.3344872

[16] Lin Wang, Hristijan Gjoreski, Mathias Ciliberto, Paula Lago, Kazuya Murao, Tsuyoshi Okita, and Daniel Roggen. 2020. Summary of the sussex-huawei locomotion-transportation recognition challenge 2020. 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* (New York, NY, USA). ACM, New York, NY, USA, 351–358. https://doi.org/10.1145/3410530.3414341

[17] 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. https://doi.org/10.1109/ACCESS.2019.2890793

[18] Lin Wang, Hristijan Gjoreskia, Kazuya Murao, Tsuyoshi Okita, and Daniel Roggen. 2018. Summary of the Sussex-Huawei Locomotion-Transportation Recognition Challenge. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (New York, NY, USA). Association for Computing Machinery, New York, NY, USA, 1521–1530. https://doi.org/10.1145/3267305.3267519

[19] Jijie Zhu. 1992. Calculation of geometric dilution of precision. *IEEE Trans. Aerospace Electron. Systems* 28, 3 (1992), 893–895. https://doi.org/10.1109/7.256323

[20] Yida Zhu, Haiyong Luo, Runze Chen, Fang Zhao, and Lin Su. 2020. DenseNetX and GRU for the sussex-huawei locomotion-transportation 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* (New York, NY, USA). ACM, New York, NY, USA, 373–377. https://doi.org/10.1145/3410530.3414349