# POIDEN: Position and Orientation Independent Deep Ensemble Network for the Classification of Locomotion and Transportation Modes

Masud Ahmed
University of Dhaka
masud.apece1314@gmail.com

Anindya Das Antar
University of Dhaka
anindyadasantar@gmail.com

Tahera Hossain
Kyushu Institute of Technology
tahera.moni@gmail.com

Sozo Inoue
Kyushu Institute of Technology
sozo@brain.kyutech.ac.jp

Md Atiqur Rahman Ahad
University of Dhaka
atiqahad@du.ac.bd

## Abstract

Sensor-based recognition of locomotion and transportation modes has numerous application domains including urban traffic monitoring, transportation planning, and healthcare. However, the use of a smartphone in a fixed position and orientation in previous research works limited the user behavior a lot. Besides, the performance of naive methods for position-independent cases was not up to the mark. In this research, we have designed a position and orientation independent deep ensemble network (POIDEN) to classify eight modes of locomotion and transportation activities. The proposed POIDEN architecture is constructed of a Recurrent Neural Network (RNN) with LSTM that is assigned the task of selecting optimum general classifiers (random forest, decision tree, gradient boosting, etc.) to classify the activity labels. We have trained the RNN architecture using an intermediate feature set (IFS), whereas, the general classifiers have been trained using a statistical classifier feature set (SCFS). The choice of a classifier by RNN is dependent upon the highest probability of those classifiers to recognize particular activity samples. We have also utilized the rotation of acceleration and magnetometer values from phone coordinate to earth coordinate, proposed jerk feature, and position insensitive features along with parameter adjustment to make the POIDEN architecture position and orientation independent. Our team "Gradient Descent" has presented this work for the "Sussex-Huawei Locomotion-Transportation (SHL) recognition challenge".

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

*UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom*
© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-6869-8/19/09…$15.00
https://doi.org/10.1145/3341162.3345570

***Keywords*** Deep learning; ensemble classifiers; neural networks; orientation independent; position insensitive; statistical features; SHL recognition challenge

**ACM Reference Format:**
Masud Ahmed, Anindya Das Antar, Tahera Hossain, Sozo Inoue, and Md Atiqur Rahman Ahad. 2019. POIDEN: Position and Orientation Independent Deep Ensemble Network for the Classification of Locomotion and Transportation Modes. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers (UbiComp/ISWC '19 Adjunct), September 9–13, 2019, London, United Kingdom.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3341162.3345570

## 1 Introduction

Due to the elevation of technological advancement in smartphones along with other benefits like user comfort, no security concern, and daily usage, embedded sensing modules have elevated the research field to discover the knowledge of user activities and transportation modes. Context-aware applications like parking spot detection, traffic route monitoring [9], health and activity monitoring [4], useful content delivery, optimizing smartphone operation in transports, etc. can be accomplished by obtaining smartphone sensor data during locomotion and traveling using transportation modes. Besides, research on transportation modes can help in road condition analysis, probabilistic mobility and locomotion mode creation, designing novel localization techniques, and so on [3]. Due to the limitation of datasets to analyze locomotion and transportation modes, the Sussex-Huawei Locomotion- Transportation (SHL) dataset [3], [12] with eight modes of locomotion and transportation in real life settings can play an important role in this research field. Sussex-Huawei Locomotion-Transportation (SHL) recognition challenge, 2019 is intended to recognize those eight modes of activities from inertial sensor data by keeping the smartphone in different body positions. The primary challenge of the participants is to design a phone position and orientation-independent model.

In most of the previous research works, the smartphones were kept in a fixed position (trouser pocket or in hand) to maintain better recognition performances. In real life cases, smartphones can be carried in different orientations, while most of the earlier works fixed the smartphone orientation, which can cause overfitting in performance [10]. However, dependency on a fixed position and orientation limit user behavior a lot and this is the reason to build orientation and position-independent model.

Research work [5] used projection-based method and position-insensitive features to handle orientation and position dependence. Research work [2] used angular acceleration to recognize position, whereas, research work [8] used light and proximity sensors to identify the position, followed by position-insensitive features. In the case of position identification, Incel et al. [6] shows 93.76% accuracy and PACP (Parameters Adjustment Corresponding to smartphone Position) [13] method shows 93.28% accuracy (91.27% accuracy for activity recognition).

In this paper, we have proposed a position and orientation independent deep ensemble network (POIDEN) to classify the eight modes of locomotion and transportation activities using the Sussex-Huawei Locomotion- Transportation (SHL) dataset. As this dataset contains bag, hips, and torso data in the training set and hand data in the test set, this is necessary for the designed pipeline to be position independent. However, our proposed POIDEN architecture is constructed of a recurrent neural network (RNN) with LSTM and five general classifiers. The task of this RNN network is to assign classifier for the final stage of activity classification based on the performance of those classifiers to classify activities with higher probabilities. This idea helped our team named "Gradient Descent" to utilize the strength of a group of classifiers to classify different activities, whereas previous works only utilize only one classifier based on overall performance. The rest of the paper is organized as follows.

After discussing about the background in the section 1, we have given a basic description of the Sussex-Huawei Locomotion- Transportation (SHL) dataset in section 2. We have represented some methods of making our system orientation and position independent in section 3. These methods include derotation of sensor values from phone coordinate to earth coordinate system, calculation of magnitude and jerk, and parameter adjustment to make some features position invariant. Section 4 represents two feature sets: intermediate feature sets (IFS) and statistical classifier feature sets (SCFS). In section 5, we have described about our proposed deep ensemble network constructed of RNN with LSTM and five general classifiers. The task of RNN is to predict the optimum classifier for final activity prediction based on the probability estimation so that we can utilize the strength of each classifier to recognize a particular activity. In section 6, we have presented our result and analysis and finally concluded the paper in section 7.

## 2 Dataset Description

The Sussex-Huawei Locomotion- Transportation (SHL) dataset [3], [12] contains 59 days of data for training, 3 days of data for the validation, and test data of 20 days. This dataset has been prepared using inertial sensor data of a smartphone by placing it in an independent manner. The generation of train, validation, and test data follows the process of segmentation using a non-overlapping sliding window of 5s duration and 100 Hz sampling rate.

The train data contains labeled raw sensor data from one user by carrying the smartphone in three locations (bag, hips, and torso). Seven inherent smartphone sensor data including acceleration, gravity, gyroscope, linear acceleration, magnetic field, the orientation of the device in quaternions, and atmospheric pressure have been provided. Each training file contains 196072 lines × 500 columns, which corresponds to 196072 frames each containing 500 samples (5s at the sampling rate 100 Hz). The frames of the training set are consecutive in time.

The validation set contains labeled sensor data from the same user from four locations (bag, hip, torso, and hand). Each file in the validation data comprises 12177 lines × 500 columns, corresponding to 12177 frames each containing 500 samples and the frames are not consecutive in time.

The test data only contains unlabeled sensor data from hands phone. The files contain 55811 lines × 500 columns. To create a challenge in real time for the researchers, the frames in the test set are shuffled and likely not consecutive in time. This dataset is intended for the Sussex-Huawei Locomotion Challenge 2019, where the goal is to recognize 8 modes of locomotion and transportation (activities) including car, bus, train, subway, walk, run, bike, and still. The aim is to identify the user activity from handphone data, but training the model using data from smartphones on other different positions (bag, hips, torso) by designing a phone-position independent activity recognition model.

## 3 Preprocessing

### 3.1 Orientation independent representation

While using inertial sensor data for recognizing activities, data are collected relative to the smartphone coordinate system. This is why even if there is no change in activity but a change in smartphone orientation, accelerometer and magnetometer data changes. In real life cases, if users hold a smartphone in different orientations, the sensor data will change even if they are performing the same activity. In order to solve the problem of overfitting due to the specific orientation of the phone, we have taken the following approaches.

#### 3.1.1 The horizontal and vertical component of acceleration

In order to avoid the impact of device orientation, we converted the coordinate system of acceleration into vertical and horizontal components by taking inspiration from the work [13]. We had both user acceleration (linear acceleration data) and gravity (gravity sensor data) data from the SHL dataset.

The gravity vector always points to the center of the earth. This is why the user acceleration (linear acceleration) can be separated into two orthogonal vectors. One pointing to the center of the earth like gravity, which can be called vertical acceleration ($A_v$). The other is perpendicular to it and this can be called horizontal acceleration, ($A_h$). If the angle between gravity ($A_g$) and user linear acceleration ($LA_u$) is considered as $\theta$, we can calculate $A_v$ and $A_h$ using following equations,

$$cos\theta = \frac{{A_g}_x{LA_u}_x + {A_g}_y{LA_u}_y + {A_g}_z{LA_u}_z}{\sqrt{{A_g}_x^2 + {A_g}_y^2 + {A_g}_z^2}\sqrt{{LA_u}_x^2 + {LA_u}_y^2 + {LA_u}_z^2}} \tag{1}$$

$$A_v = \sqrt{{LA_u}_x^2 + {LA_u}_y^2 + {LA_u}_z^2}\, cos\theta \tag{2}$$

$$A_h = \sqrt{{LA_u}_x^2 + {LA_u}_y^2 + {LA_u}_z^2}\, sin\theta \tag{3}$$

where, $LA_u$ = user linear acceleration, $A_g$ = acceleration due to gravity, $A_v$ = vertical acceleration, and $A_h$ = horizontal acceleration

#### 3.1.2 Magnitude and jerk

In this stage, we have calculated the magnitude value of raw linear acceleration, acceleration, and magnetometer value using the following formulas,

$$A_{mag} = \sqrt{A_x^2 + A_y^2 + A_z^2} \tag{4}$$

$$M_{mag} = \sqrt{M_x^2 + M_y^2 + M_z^2} \tag{5}$$

$$LA_{mag} = \sqrt{LA_x^2 + LA_y^2 + LA_z^2} \tag{6}$$

where, $A_{mag}$, $M_{mag}$, and $LA_{mag}$ are magnitude values of accelerometer, magnetometer, and linear accelerometer

We have also calculated the jerk (rate of change) value of horizontal acceleration ($A_h$), vertical acceleration ($A_v$), and magnitude of linear acceleration using the following formulas,

$$A_h\ (jerk) = \frac{d}{dt}(A_h) \tag{7}$$

$$A_v\ (jerk) = \frac{d}{dt}(A_v) \tag{8}$$

$$LA_{mag}\ (jerk) = \frac{d}{dt}(LA_{mag}) \tag{9}$$

#### 3.1.3 Coordinate conversion using rotation matrix

We have converted raw accelerometer and magnetometer data from phone coordinate system to Earth North-East-Down coordinate system inspired by the research work [7]. This method helped to obtain orientation-independent sensor information. We have formed a rotation matrix $R_{NB}$ using quaternion values (w, x, y, z) and used the following equations for coordinate conversion.

$$R_{NB} = \begin{bmatrix} 1-2(y^2+z^2) & 2(xy-wz) & 2(xz+wy) \\ 2(xy+wz) & 1-2(x^2+z^2) & 2(yz-wx) \\ 2(xz-wy) & 2(yz+wx) & 1-2(x^2+y^2) \end{bmatrix} \tag{10}$$

$$\begin{bmatrix} A_X \\ A_Y \\ A_Z \end{bmatrix} = R_{NB} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} \tag{11}$$

$$\begin{bmatrix} Mag_X \\ Mag_Y \\ Mag_Z \end{bmatrix} = R_{NB} \begin{bmatrix} Mag_x \\ Mag_y \\ Mag_z \end{bmatrix} \tag{12}$$

where, $A_x, A_y, A_z$ and $Mag_x, Mag_y, Mag_z$ are raw acceleration and magnetometer values, $A_X, A_Y, A_Z$, and $Mag_X, Mag_Y, Mag_Z$ are converted values, $R_{NB}$ is the rotation matrix, and $Ori = (w, x, y, z)$ is orientation of the device in quaternions

As we have access to both accelerometer and linear accelerometer data along with gravity value, we have calculated horizontal and vertical components of acceleration earlier using linear acceleration and gravity data, where vertical acceleration follows similar pattern irrespective of smartphone positions in body. In this section, we have also utilized this coordinate conversion approach, which additionally helps us to utilize raw acceleration value for orientation invariant representation.

### 3.2 Position independent representation

As we mentioned earlier if the smartphones are kept in a specific position while data collection, the trained model gets biased to that specific position to show better performance. In order to deal with the real-life challenges, we have intended to design a position independent model.

In order to make the model position independent, we have used three methods. First of all, we have omitted those sensor-based features that are very much sensitive to the position like a gyroscope. Secondly, we have chosen position insensitive features. Thirdly, we have adjusted minimum position sensitive features by a value, which reflects the intensity of activities in a specific position.

We can see the four graphs of walking data from four different positions reflecting horizontal acceleration, vertical

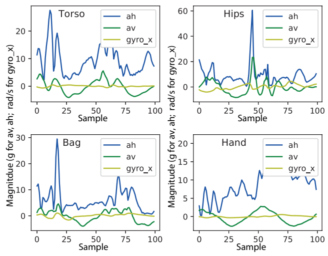


**Figure 1.** Visualization of sensor data from four body positions for walking activity.

acceleration, and gyroscope data in Figure 1. We can see that waveforms of the gyroscope are different in four positions. The change of gyroscope value is small in hand than other positions as users normally make the phone stationary relative to their body for texting while walking or swing it along with the arm while keeping it in hand. In the hips, the change rate of the gyroscope value is high, which means gyroscope value is very much position sensitive. Horizontal acceleration has minimal changes in four positions as this has a higher value in hand and hips than bag and torso because of the movement of users hand and leg. We can see that vertical acceleration keeps almost similar change in four positions for a specific activity that makes it position insensitive and reflects the characteristics of motion. Besides, we intended to adjust horizontal acceleration to better the performance to minimize the differences in four position. We have selected a parameter that reflects the intensity of those activities in different positions or that can reflect the position information best. As we mentioned earlier, gyroscope data has the highest rate of change in four positions for the same activity, we have taken the average of the maximum change rate of the gyroscope value along three axes and named it $\beta$. It is also mentioned in [13] that gyroscope data has the highest dissimilarity in different positions. This $\beta$ parameter can exhibit the intensity degree of movement in different positions. We have taken four $\beta$ for four different positions ($\beta_{hand}$, $\beta_{bag}$, $\beta_{hips}$, and $\beta_{torso}$) and used the following equation to adjust horizontal acceleration.

$$A_{horizontal(adjusted)} = \frac{\beta + 1}{\beta} A_{horizontal} \tag{13}$$

where, $A_{horizontal(adjusted)}$ is the adjusted horizontal acceleration, $A_{horizontal}$ is horizontal acceleration, and $\beta$ is the average of the maximum change rate of the gyroscope value along three axes for each position

## 4 Features

We have considered two sets of feature (Intermediate Feature Set (IFS) and Statistical Classifier Feature Set (SCFS)) that are orientation and position insensitive. IFS is the time series modified signals that have been extracted in the preprocessing stage and the size of this set is 500 samples × 12 channels

**Table 1.** Intermediate Feature Sets (IFS) with time information for training RNN with LSTM.

| Sensor Name | Intermediate feature set (IFS) |
|---|---|
| LAcc, Gra | Vertical Acc ($A_v$) |
| LAcc, Grav and Gyr | Position insensitive Horizontal Acc ($A_h$) |
| LAcc | Magnitude of LAcc |
| Ori and Acc | Derotated Acc (x, y, z) |
| Acc | Magnitude of Acc |
| Ori and Mag | Derotated Mag (x, y, z) |
| Mag | Magnitude of Mag |
| Pressure | Pressure |

Lacc: Liner Accelerometer, Gra = Gravity,
Gyr = Gyroscope, Ori = Orientation,
Acc: Accelerometer, Mag = Magnetometer

**Table 2.** Statistical Classifier Feature Sets (SCFS) extracted from intermediate feature set to train general classifiers.

| Intermediate feature set | Domain | Statistical classifier feature set (SCFS) |
|---|---|---|
| Derotated Acc (x, y, z)<br>Derotated Mag (x, y, z)<br>Magnitude of Acc<br>Magnitude of Mag | Time | Mean, Std, Var<br>Max, Min, MAD |
| | Frequency | Mean freq, Kurtosis<br>Skew, Energy, Entropy |
| $A_v$ (Jerk)<br>$A_h$ (Jerk)<br>Magnitude of Acc (Jerk)<br>Magnitude of Lacc (Jerk) | Time | Mean, Std, Var<br>Max, Min, MAD |
| | Frequency | Mean freq, Kurtosis<br>Skew, Energy, Entropy |
| $A_v$ | Time | Range, Max value of differentiation |
| $A_h$ | Time | Mean |
| Magnitude of Lacc | Time | Mean, Var, Range, Velocity |
| | Frequency | Energy |
| Pressure | Time | Mean |

for each window of 5s duration. SCFS is a statistical aspects of IFS in both time and frequency domain. For each window, SCFS is a one-dimensional vector and we have calculated 140 feature vector in total under SCFS. As, IFS features contain time series information of sensor data, this has been used to train the neural network (RNN with LSTM), whereas SCFS has been used to train the ensemble of general classifiers for final activity recognition. In Table 1 and Table 2, we have summarized the features. While choosing statistical features, we have extracted the features both in time and frequency domain (250 bins) following our previous work [1]. The main challenge, in this case, was to select the features based on the sensor modules to express the difference between transportation modes. For example, we have found out from our experiment that, the air pressure while traveling with buses has a smaller mean value than the mean pressure while traveling with cars. Similarly, the mean pressure is larger in the subway than train and in different body positions, the air pressure values almost follow a similar pattern. This is the reason behind choosing mean pressure under SCFS. We have performed similar experiments to choose other important feature sets.

We mentioned earlier the training, validation, and test set contain segments of 5s frames. As device frequency was 100 Hz, each frame contains 500 samples. To extract statistical features (SCFS), we have used these segmented frames of

5s duration containing 500 samples. The choice of not segmenting the 5s window further is the reason that in real life it is not possible to switch transportation modes within 5s. So, our method follows that within 5s there can not be any possible transitional activities. The labeled training and validation set also resemble these characteristics.

# 5 Methodology

## 5.1 Deep Ensemble Network

In this research, we have designed a deep network comprises of a Recurrent Neural Network (RNN) with Long short-term memory (LSTM) and an ensemble of five general classifiers (Random Forest (RnF), Linear Discriminant Analysis (LDA), Logistic Regression (LR), Decision Tree (DT), and Gradient Boosting (GradBoost)). We have found out from experiments that, a single classifier is not efficient enough to classify all the locomotion and transportation activities with higher accuracy and precision. It may happen that, a specific activity can be better classified by one specific classifier but in case of classifying other activities, that classifier shows poor performance. Previous research works solved this problem by choosing a classifier that shows the overall best performance. In this research, we have tried to ensemble a group of classifier so that in the classifying stage, based on the input test sample characteristics, the optimum classifier will be chosen. If the input sample changes, the classifier will change at the classification stage. In order to choose the optimum classifier with the best performance for an incoming activity sample, we have trained an RNN with LSTM. The task of this deep network is to assign the best classifier at the classification stage based on the input sample. In the training stage, the RNN is trained in such a way that it remembers the classifier (for example, random forest) that has classified an incoming activity sample (for example, walk) with the highest percentage of probability. If random forest predicts walk with the highest amount of probability than other classifiers for most of the samples, the RNN will assign random forest classifier in the test stage whenever RNN will show a walk alike sample in the input. Besides, as we have used some methods to make our features orientation and position independent, we have named our network as position and orientation independent deep ensemble network (POIDEN).

In Figure 2, we have shown the series of processing steps of our deep ensemble network (POIDEN) in case of the training phase. The training phase has been divided into two segments. First, we train five general classifiers (Random forest with 30 trees, decision tree, Gradient boosting for 50 boosting stages in a forward stage-wise fashion, Linear discriminant analysis, and Logistic Regression) using our statistical classifier feature set (SCFS) for the training data and calculated the probability of correctly classifying a particular activity for each classifier by using the validation data. After evaluation of these classifiers, we have seen that different classifier predicts different types of activity with high accuracy and precision but overall accuracy of a single classifier is not good enough. Therefore, we have designed an RNN with LSTM, which predicts a single general classifier among the five for each activity, that can recognize a specific type of activity signal more accurately than others. So, the final model is the RNN with LSTM trained by IFS, which is trained to provide the optimum general classifier name among the mentioned five in the output to predict final activity. This RNN model with ensemble of classifiers has been used to classify the test data submitted to the challenge.

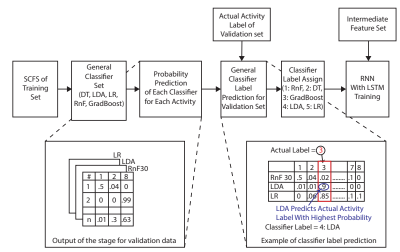


**Figure 2.** Series of processing steps of our deep ensemble network (POIDEN) in training phase.

The testing stage of this network has been shown in Figure 3. In the test stage, the intermediate feature set (IFS) for the test data is fed into RNN with LSTM network. The network predicts the optimum classifier (for example LDA) for each input sample and for that input sample that optimum classifier (LDA) recognizes the final activity label. In order to recognize the activity label, the optimum classifier uses the statistical classifier feature set (SCFS) of the same input test sample fed to RNN.

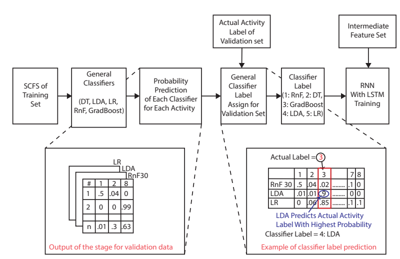


**Figure 3.** Basic architecture of our deep ensemble network (POIDEN) in testing phase.

We have designed the model RNN with two stacked LSTM cells. The number of Hidden layers is 32. We have used a total iteration number 2,000,000 and for each iteration, the batch size is 1500. We have used ReLu as the activation function.

## 5.2 Data preparation and computational resources

We have used three sets of data for this process. First of all, we have used 35% data from the validation set as our validation set. We have added the rest of the data with training data

**Table 3.** Confusion matrix of the validation set on hand position data only.

| Activity | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1 | **456** | 11 | 0 | 38 | 24 | 26 | 27 | 21 |
| 2 | 7 | **387** | 3 | 158 | 0 | 1 | 0 | 11 |
| 3 | 0 | 1 | **86** | 2 | 0 | 0 | 0 | 0 |
| 4 | 21 | 2 | 0 | **318** | 3 | 9 | 1 | 2 |
| 5 | 23 | 5 | 0 | 25 | **535** | 78 | 23 | 24 |
| 6 | 54 | 10 | 1 | 60 | 118 | **396** | 25 | 14 |
| 7 | 33 | 6 | 0 | 6 | 30 | 23 | **312** | 233 |
| 8 | 26 | 4 | 0 | 11 | 38 | 13 | 135 | **386** |

1: Still, 2: Walk, 3: Run, 4: Bike, 5: Car, 6: Bus, 7: Train, 8: Subway

and used as our training set. Training set and validation set both contain the activity label file. We did not modify the provided test set. As the test set does not contain the label file, we have used the validation set for the evaluation of our method.

In this research, we have used CPU processor i5 with quad cores (Clock speed: 2.5GHz, RAM: 8GB). The training time for the models are Random Forest: 14m, Decision Tree: 8m, Logistic Regression: 1h 31m, Gradient Boosting: 2h 38m, and RNN: 2h 10m (Total: 6h 41m). It required 9m and 16s to evaluate the test set. The final trained model size is 383.1 MB.

## 6 Result and Analysis

We have shown the performance of our method in Table 3. As, we have to deal with the challenge of building position and orientation independent model, our method shows an overall F1 score of 70.57% (still: 74.57%, walk: 77.95%, run: 96.09%, bike: 65.3%, car: 73.24%, bus: 64.71%, train: 53.52%, subway: 59.20%) and overall accuracy is 67.5%. The most challenging part was to classify between train and subway as they exhibit similar pattern in terms of acceleration and velocity. Without train and subway, the overall accuracy rises up to 76.20%. User sitting behavior also remains same in the car, bus, train, and subway. This can be another reason of misclassification among these activities.

## 7 Conclusion

In this paper, we have mainly focused to design a position and orientation independent deep ensemble network (POIDEN) as we have to train our network from the body, hips, and torso sensor data, whereas the test set contains hand sensor data in the case of SHL dataset. We have utilized axis transformation of acceleration and magnetometer data into earth coordinate system along with magnitude and jerk-based feature to deal with the orientation challenges. Besides, we have adjusted horizontal acceleration by the average maximum change rate of gyroscope value to make this position independent along with some other position independent features. Finally, we have trained a deep recurrent neural network to predict optimum statistical classifiers (random forest, decision tree, logistic regression, etc.) for the final activity label prediction. However, our method can be improved by adding features that can classify train and subway well as they are very much similar in nature. The recognition result for the testing dataset will be presented in the summary paper of the challenge [11].

## References

[1] Anindya Das Antar, Masud Ahmed, Mohammad Shadman Ishrak, and Md Atiqur Rahman Ahad. 2018. A comparative approach to classification of locomotion and transportation modes using smartphone sensor data. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. ACM, 1497–1502.

[2] Doruk Coskun, Ozlem Durmaz Incel, and Atay Ozgovde. 2015. Phone position/placement detection using accelerometer: Impact on activity recognition. In *2015 IEEE Tenth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*. IEEE, 1–6.

[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.

[4] Hristijan Gjoreski, Boštjan Kaluža, Matjaž Gams, Radoje Milić, and Mitja Luštrek. 2015. Context-based ensemble method for human energy expenditure estimation. *Applied Soft Computing* 37 (2015), 960–970.

[5] Apiwat Henpraserttae, Surapa Thiemjarus, and Sanparith Marukatat. 2011. Accurate activity recognition using a mobile phone regardless of device orientation and location. In *2011 International Conference on Body Sensor Networks*. IEEE, 41–46.

[6] Ozlem Incel. 2015. Analysis of movement, orientation and rotation-based sensing for phone placement recognition. *Sensors* 15, 10 (2015), 25474–25506.

[7] Vito Janko, Nina Reščič, Miha Mlakar, Vid Drobnič, Matjaž Gams, Gašper Slapničar, Martin Gjoreski, Jani Bizjak, Matej Marinko, and Mitja Luštrek. 2018. A New Frontier for Activity Recognition: The Sussex-Huawei Locomotion Challenge. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. ACM, 1511–1520.

[8] Fen Miao, Yi He, Jinlei Liu, Ye Li, and Idowu Ayoola. 2015. Identifying typical physical activity on smartphone with varying positions and orientations. *Biomedical engineering online* 14, 1 (2015), 32.

[9] Joshua E Siegel, Dylan C Erb, and Sanjay E Sarma. 2017. A survey of the connected vehicle landscape—Architectures, enabling technologies, applications, and development areas. *IEEE Transactions on Intelligent Transportation Systems* 19, 8 (2017), 2391–2406.

[10] Yunus Emre Ustev, Ozlem Durmaz Incel, and Cem Ersoy. 2013. User, device and orientation independent human activity recognition on mobile phones: Challenges and a proposal. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, 1427–1436.

[11] 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 *Proc. HASCA 2019*.

[12] 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.

[13] Rong Yang and Baowei Wang. 2016. Pacp: A position-independent activity recognition method using smartphone sensors. *Information* 7, 4 (2016), 72.