




# A deep semi-supervised machine learning algorithm for detecting transportation modes based on GPS tracking data

Paria Sadeghian<sup>1</sup>  · Arman Golshan<sup>1</sup> · Mia Xiaoyun Zhao<sup>1</sup> · Johan Håkansson<sup>1</sup>

Accepted: 6 February 2024  
© The Author(s) 2024

## Abstract

Transportation research has benefited from GPS tracking devices since a higher volume of data can be acquired. Trip information such as travel speed, time, and most visited locations can be easily extracted from raw GPS tracking data. However, transportation modes cannot be extracted directly and require more complex analytical processes. Common approaches for detecting travel modes heavily depend on manual labelling of trajectories with accurate trip information, which is inefficient in many aspects. This paper proposes a method of semi-supervised machine learning by using minimal labelled data. The method can accept GPS trajectory with adjustable length and extract latent information with long short-term memory (LSTM) Autoencoder. The method adopts a deep neural network architecture with three hidden layers to map the latent information to detect transportation mode. The proposed method is assessed by applying it to the case study where an accuracy of 93.94% can be achieved, which significantly outperforms similar studies.

**Keywords** Travel identification · LSTM Autoencoder · Unsupervised learning · Deep learning · GPS tracking data

## Introduction

Understanding travel and human mobility behaviors, traffic demands, and the impact of transportation infrastructure on people is crucial in transportation science. The traditional data collection methods were physical interviews, paper-based travel diaries, travel surveys, phone interviews, and internet web pages, where volunteers report why and how they travel to a destination (Shen and Stopher 2014; Wu et al. 2016). However, the traditional methods are time-consuming and error-prone, resulting in a low response rate (Dutta and Patra 2023; Tamim Kashifi et al. 2022).

---

✉ Paria Sadeghian  
psd@du.se

<sup>1</sup> School of Technology and Business Studies, Dalarna University, Falun SE-79188, Sweden

Global Positioning System (GPS) devices are ubiquitous tools used to record spatial and temporal information of human movement during trips and have significant application in transportation sector (Raza et al. 2022; Zhao et al. 2017). However, GPS tracking data can only record positional recording characteristics such as latitude, longitude, altitude, speed, and time without any information regarding which transportation mode is used during a trip. Therefore, using the large GPS tracking dataset to understand human behaviour patterns in urban areas has led to various behavioral applications, including frequently visited locations, transportation mode detection, location-based activity (Ma et al. 2023; Roy et al. 2022). Specifically, transportation mode detection studies, several data processing methods, and machine learning algorithms have been applied (Dabiri et al. 2019; Namdarpour et al. 2022; Yu 2020). In most of the current studies, five common transportation modes are detected (walk, bike, bus, car, and train). Notably, although applying a large GPS dataset can improve the accuracy level, the structure and size of the dataset that is used in current studies varied (Li et al. 2023; Ma et al. 2023). Some studies considered only a few users and trips, while others considered more than 100 users and recorded the data for a year (Dutta and Patra 2023; Marra et al. 2022). However, it is important to note that the definition of 'large dataset' can vary depending on the specific research context and objectives.

Most of the studies in detecting transportation modes by using GPS tracking data have two steps. The first step is cleaning and processing the raw GPS tracking and removing the outliers. Each trip's features such as velocity, acceleration, heading change rate, trip distance are computed in this step. In the second step, all features are applied to an algorithm to detect the transportation modes. Three main methodological approaches are commonly used in this step to detect the transportation modes: machine learning algorithms, statistical methods, and rule-based algorithms. Among these methods, machine learning methods, including supervised learning algorithms, semi-supervised learning, and unsupervised learning algorithms, are three approaches. A wide range of traditional machine learning algorithms such as decision tree, random forest, support vector machine has been widely applied (Nitsche et al. 2014; Sadeghian et al. 2022).

Fully supervised learning algorithms are most commonly used, and they require identification and formulation of features manually before applying a machine learning algorithm, in order to detect the transportation modes. Unsupervised learning algorithms used a fully unlabeled GPS tracking dataset to detect the correct transportation mode (Lin et al. 2013; Patterson et al. 2003; Weinstein et al. 2010; Yazdizadeh et al. 2019).

Using semi-supervised learning can improve the efficiency while the accuracy can be retained even increased since only a small portion of raw GPS data needs to be labeled. Recently, a few studies have used the semi-supervised learning algorithm to detect transportation modes (Dabiri et al. 2019; Dutta and Patra 2023; Markos and Yu 2020; Zhang et al. 2022). However, these approaches rely on the relatively old GPS tracking dataset, and the accuracy level of the proposed methods is rather low (below 80%).

This study proposes a novel semi-supervised deep learning algorithm to explore further the advantages of applying semi-supervised learning methods. This algorithm is applied to a large volume GPS tracking dataset. The potential for LSTM Autoencoder, and deep neural network (DNN) is examined. This attempt exposes the possibility of applying a more straightforward and efficient model to detect transportation modes.

The rest of this paper is organized as follows. Section 2 introduces the literature review of transportation mode detection research. Section 3 presents the proposed method in detail.

Section 4 presents the empirical analysis by applying the algorithm to accurate GPS data and discusses the results. Section 5 concludes the paper with highlights and main takeaways.

## Literature review

To date, a large part of the literature has proposed methodology for transportation mode detection based on different data sources, including raw GPS tracking data (Dutta and Patra 2023; Tamim Kashifi et al. 2022; Xiao et al. 2017), accelerometers (Alam et al. 2023; Nick et al. 2010), mobile-phone networks (Bachir et al. 2019; Xu et al. 2022; Zhagyparova et al. 2023). A wide range of methods, such as rule-based methods, statistical methods, as well as the traditional supervised learning algorithms have been employed for detecting transportation modes, including, random forest, decision tree, support vector machine, and multi-layer perceptron (Li et al. 2021; Sadeghian et al. 2021; Wu et al. 2016). Using supervised learning algorithms to detect transportation mode requires the GPS tracking dataset to be fully labeled (Giri et al. 2022; Tamim Kashifi et al. 2022). Some studies have integrated Geographical information layers with GPS tracking data to improve the method's performance (Gong et al. 2012; Lee and Kwan 2018; Li et al. 2021; Vinayaraj and Mede 2022). However, these studies used only one type of sensor to practically reach more accurate results (Xiao et al. 2017). This section focuses on reviewing studies that have used GPS tracking data to detect transportation modes using machine learning methods. Table 1. Shows the summary of the methods in transport mode detection studies.

Using a fully labelled dataset has been found challenging because manually labelling the data is time-consuming and contains human errors (Dabiri et al. 2020; Reddy et al. 2010; Roy et al. 2022). To improve classification efficiency, some transportation mode detection studies have employed GIS layer information into labeling (Roy et al. 2022). Roy et al. (2022) study, focuses on enhancing transportation mode detection using GPS data by incorporating geographic context. Results revealed that adding geographic context significantly improved accuracy, particularly with context-specific models, emphasizing the importance of local geography in mode detection. In other study done by Sadeghian et al. (2022) proposed a stepwise methodology by combing three common methods (unsupervised learning method, GIS multi criteria-process, and supervised learning algorithms) that can achieve a high accuracy of 99% with use of only 10% labelled data. In our previous work (Sadeghian et al. 2022), we reported a 99% transport mode detection accuracy with the use of a different methodology, which was applicable to a relatively small dataset. In that case, traditional machine learning methods, which require less data, were feasible. However, in the current study, we are dealing with a much larger dataset, which makes it impractical to apply traditional supervised learning methods due to their data-intensive nature. This shift in dataset size and complexity necessitated the development of the current semi-supervised deep learning approach, which can effectively harness unlabeled data to enhance its accuracy and generalizability. Therefore, the choice of methodology is highly dependent on the dataset's characteristics and scale, with the current approach being tailored to the challenges posed by larger and more diverse datasets.

Few approaches have used unsupervised learning to detect the transportation mode using fully unlabeled GPS tracking data. Patterson et al. (2003) proposed an unsupervised learning algorithm to detect three transportation modes (walk, car, and bus). This study used

**Table 1** Summary of methods in transportation mode detection studies

Authors	Methods	Dataset	Transportation modes	Users	Accuracy (%)
Patterson et al. (2003)	Unsupervised	GPS device	Walk, Car, Bus	1	84
Zheng et al. (2008b)	Supervised	GPS device	Walk, Bike, Car, Bus	45	90
Lin et al. (2013)	Unsupervised	GPS device	Walk, Bike, Car, Bus	101	76
Lari and Golroo (2015)	Supervised	GPS device	Walk, Car, Bus	35	96
Feng and Timmermans (2016)	Supervised	GPS device	Stationary, Walk, Running, Bike, Motor, Car, Bus, Train, Metro	8	99
Rezaie et al. (2017)	Semi-Supervised	Mobile phone app	Walk, Bike, Car, Bus, Train	796	80
Xiao et al. (2017)	Ensemble methods	GPS device	Walk, Bike, Car, Bus, Metro, Train	182	90
Dabiri and Heaslip (2018)	Unsupervised	GPS device	Walk, Bike, Car, Bus, Train	182	84
Yazdizadeh et al. (2019)	Unsupervised	Mobile phone app	Walk, Bike, Car, Bus	NA	91
Yu (2020)	Semi-Supervised	GPS device	Walk, Bike, Car, Bus, Train	182	91
Li et al. (2021b)	Entropy-based Encoder-Decoder	GPS device	Walk, Bike, Car, Bus, Train	182	79
Namdarpour et al. (2022)	Genetic programming	GPS device	Walk, Bike, Car, Bus, Train	182	83
Roy et al. (2022)	Supervised	GPS device	Walk, Bike, Car, Bus, Train	NA	60
Sadeghian et al. (2022)	Supervised	GPS device	Walk, Bike, Car, Bus, Train	20	99
Dutta and Patra (2023)	Unsupervised	GPS device	Walk, Bike, Car, Bus, Train	182	70

an Expectation-Maximization (EM) method and speed as a crucial feature to distinguish between different modes. Moreover, in this study, the GIS layer information such as a bus-road network, bus station, and parking spots was used. The model detected the correct transportation mode with an accuracy of 84%, for 29 trips of one individual. In another study, Lin et al. (2013) proposed an unsupervised algorithm (MO detect algorithm) based on the Kolmogorov-Smirnov test. In this study, the model detected four-transportation modes (walk, bike, car, and bus) with an accuracy of 76%. The features computed in this study were average speed and maximum speed. However, the dataset used by Lin et al. (2013) was relatively small and for only one volunteer over ten months. Dutta and Patra (2023) addressed the challenge of transportation mode detection using unsupervised learning due to limited labeled data and a large amount of unlabeled data. They proposed a method that combines point-level characteristics with GPS coordinates to extract joint probability densities using masked autoregressive flow (MAF) and applied K-means to find transportation modes, demonstrating its effectiveness compared to traditional machine learning approaches.

Semi-supervised algorithms have the potential to increase and improve the performance and accuracy of transportation mode detection, compared to traditional supervised and unsupervised algorithms (Li et al. 2021b). Semi-supervised algorithms require a downsized labeled dataset, while accuracy and efficiency can be improved, compared to supervised algorithms (Kumar et al. 2021; Moreau et al. 2021). Semi-supervised learning algorithms follow two categories to detect the correct transportation modes. The first group of semi-supervised methods works with a combination of using an unsupervised learning algorithm as a pre-training step. Then, supervised learning algorithms are adjusted using the labelled dataset. The second group of semi-supervised methods works with a collaborative process where both supervised and unsupervised algorithms are trained simultaneously. A small portion of the labeled dataset is used for training, and the other unlabeled dataset is used for learning algorithms (Namdarpour et al. 2022; Yao et al. 2023).

Li et al. (2021b) presented a novel similarity entropy-based encoder-decoder (SEED) model for efficient transportation mode detection using GPS data. SEED utilized a semi-supervised learning module based on similarity entropy to enhance performance. It achieved significant improvements over baseline methods, with a 5% increase in metrics like intersection over union. Rezaie et al. (2017) proposed a semi-supervised algorithm with an accuracy of 80% in detecting five transportation modes (walk, bike, car, bus, and train). The features that were used in this study were speed, duration of the trips, length of the trips, and the start and endpoint of the trips to the transit network. In another study by Dabiri et al. (2019b), a model with a combination of neural network and autoencoder was proposed. This study used four features including speed, acceleration, the distance between two points, and Jerk, to detect five transportation modes. Yazdizadeh et al. (2019) presented a semi-supervised algorithm for transportation mode detection based on a generative adversarial network. This study used labeled and unlabeled datasets from random noise to train a generative model. However, the paper's essential technical details were not well discussed, so assessing its validity was difficult. Yu (2020) introduced semi-supervised deep ensemble learning methods to use a minimal number of labelled data to detect the five transportation modes by using the same dataset as Dabiri and Heaslip (2018) and Zheng et al. (2008). The results indicated better performance of the model compared to the other two studies in terms of accuracy. They proved that the accuracy of the collected data had no impact on the proposed model's performance. It is worth noting that the highest level of accuracy of

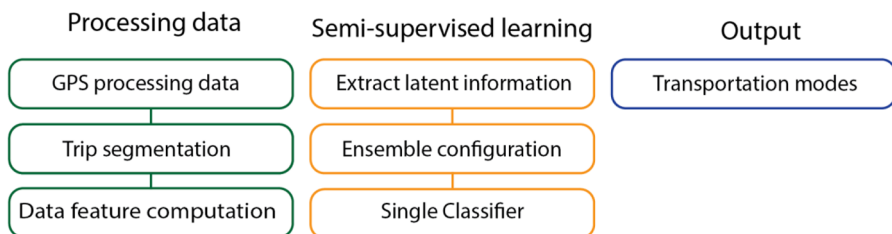
both deep learning-based studies (Dabiri and Heaslip 2018; Yu 2020) is still lower than the supervised studies (Feng and Timmermans 2016; Lari and Golroo 2015), which are purely based on the fully labelled dataset. In summary, the machine learning approach is the most common method used to detect transportation modes. Among the three machine learning methods, semi-supervised learning is the most commonly used to identify the correct transportation modes. This may be due to the fact that most of the GPS data is unlabeled. Semi-supervised algorithms work with a small amount of labelled data, while supervised learning needs the data to be fully labeled, which is time-consuming. In recent studies, deep learning has been used to detect transportation modes. To handle these disadvantages, this study in the following sections proposes a semi-supervised method to detect transportation modes based on GPS tracking dataset.

## Methodology

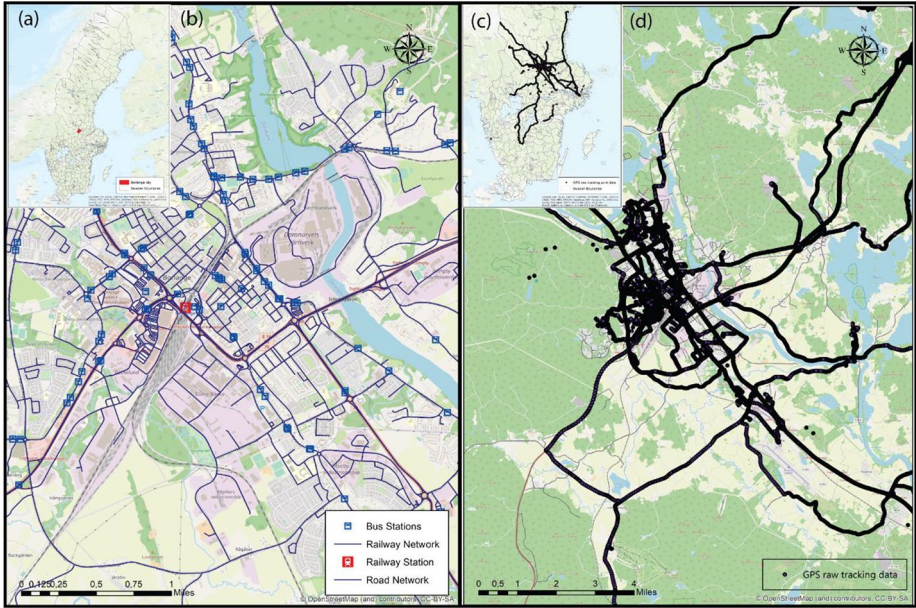
This section proposes a new semi-supervised method (Fig. 1). First, we present in detail used data pre-processing methods, and then the architecture of the proposed method is presented. Afterwards, the data feature computation and extraction of latent information schemes are created with brief introductions to related preliminaries.

### GPS dataset processing

The raw GPS tracking data were collected from September 2019 to September 2020, from 91 volunteers in Borlänge, a city located in the central part of Sweden (Fig. 2(a, b, c, d)). Each volunteer carried a portable device, “Renkforce GPS logger,” for three weeks. The device recorded date, time, latitude, longitude, elevation, and speed every five seconds. Volunteers were requested to carry the device when they started the day and continue to the end of the day. In total, around five million positional recordings were recorded, and over 6% of the dataset was labelled with the correct used transportation modes. Processing of cleaning outliers, missing values, and signal errors were conducted to prepare the dataset. All point data with a speed of more than 300 km per hour were removed from the dataset. Also, all the points that had greater timestamps than the next point were considered signal errors and were removed. If the device was stable for more than 20 min, all positional recordings before the last timestamp were regarded as one trip. In total, 11,539 trips were extracted during the time period.



**Fig. 1** The framework of the proposed method



**Fig. 2** (a, b, c, d). location of Borlänge city in Sweden (a), The public transportation network in study area (b), The raw GPS tracking data (c), The GPS tracking data in the study area

A trip may include different types of transportation modes and can be divided into a single segment, where each segment indicates a specific transportation mode. For example, a trip may start with a bus and end with walking. In this paper, we employ the Pruned Exact Linear Time (PELT) algorithm to effectively identify change points within a dataset where statistical properties change. The PELT method is rooted in an algorithm initially proposed by Jackson et al. (2005) It incorporates a key pruning step within the dynamic programming framework, as outlined by Frappart and Bourrel (2018). PELT stands out for its ability to accurately pinpoint change points and segment datasets with a higher level of precision compared to traditional binary segmentation methods, as demonstrated by Killick et al. (2012). The PELT method achieves high precision by optimizing the partitioning process and integrating it with a pruning technique, resulting in an exact and computationally efficient solution denoted as  $F(n)$ .  $F(n)$  exhibits linear computational complexity with respect to the number of data points, as expressed in Eq. 1.

$$F(n) = \min_{\tau_m} \left\{ \sum_{i=1}^{i=m+1} [C(y_{\tau_{i-1}+1} : y_{\tau_i}) + \beta] \right\} \quad (1)$$

$$F(n) = \min_{\tau_m} \left\{ \sum_{i=1}^{i=m+1} [C(y_{\tau_{i-1}+1} : y_{\tau_i}) + \beta] + C(y_{\tau_{m+1}} : y_n) \right\} \quad (2)$$

$$F(n) = \min_{\tau_m} \{ F(\tau_m) + C(y_{\tau_{m+1}} : y_n) + \beta \}. \quad (3)$$

In Eq. 1,  $F(n)$  denotes the minimization for data  $y_{1:n}$ ,  $n$  is the number of point data,  $\beta$  is the penalty to control the overfitting, and  $C$  is a cost function for the  $i^{th}$  segment.  $m$  is all number of changes in their positions,  $\tau$ .

Since PELT aims to find the optimal number of change points, the model needs to calculate the optimal number of breakpoints or optimal segmentation until that changes the minimization of  $F(n)$ . This iterative nature can lead to an inner minimization, which is shown by  $F(\tau_m)$  in Eqs. 2 and 3.

### Data feature computation

Features such as time, date, latitude, longitude, speed elevation of each positional recording can be directly extracted from GPS devices. However, in order to detect the transportation modes, more features are needed. Therefore, for each segment, the following features are calculated: the distance between two points, total distance, bearing rate, turning change rate, the time difference between two points, total duration, and features related to speed including average speed, minimum and maximum speed, acceleration, and Jerk. The features selected for our transportation mode detection model were chosen based on their established relevance in capturing essential aspects of movement patterns, as emphasized in key studies (Dabiri et al. 2020; Dutta and Patra 2023; Li et al. 2021; Markos and Yu 2020; Sadeghian et al. 2022). These studies collectively emphasized the significance of features such as distance metrics, bearing rate, turning change rate, time-related features, and speed-related metrics in accurately characterizing transportation modes. The inclusion of these features is grounded in a robust foundation of empirical evidence and established practices, ensuring the effectiveness of our model in discerning diverse transportation modes. In order to provide a comprehensive overview of the input data features used in our transportation mode detection model, we have included a summary table (Table 2) that outlines each feature’s name, data type, and description.

The Vincenty’s formula (Vincenty, 1975) is used for computing the distance between two points ( $P_1, P_2$ ), the time difference between  $P_1$  and  $P_2$  is noted as, the acceleration, average speed, Jerk, and turning rate are calculated according to Eqs. 4, 5, 6, 7, 8 and 9:

$$A_{p1} = \frac{S_{p2} - S_{p1}}{\Delta t} \tag{4}$$

$$S_{p1} = \frac{Vincenty(P_1, P_2)}{\Delta t} \tag{5}$$

$$AvgS_p = \sum_{i=0}^n \frac{S_{p_i}}{n} \tag{6}$$

$$J_i = \frac{a_{i+1} - a_i}{\Delta t_i} \quad 1 \leq i \leq N; J_N = 0 \tag{7}$$

$$\frac{Vincenty(lat_i, long_i, lat_{i+1}, long_i)}{Vincenty(lat_i, long_i, lat_i, long_{i+1})} \tag{8}$$



**Table 2** Summary of input data features for transportation mode detection model

Variable Name	Type	Description
Speed	Float	Speed of the GPS points in km per hour.
Average Speed	Float	Average speed within a segment in km per hour.
Minimum Speed	Float	Minimum speed within a segment in km per hour.
Maximum Speed	Float	Maximum speed within a segment in km per hour.
Acceleration	Float	Rate of change of speed ( $\frac{m}{s^2}$ )
Jerk	Float	Rate of change of acceleration ( $\frac{m}{s^3}$ )
Distance	Float	Distance between two consecutive GPS points
Total Distance	Float	Total distance covered within a segment in meters.
Bearing Rate	Float	Rate of change in bearing (direction) between points.
Turning Change Rate	Float	Rate of change in turning angle between points.
Time Difference	Float	Time duration between consecutive GPS data points in seconds.
Total Duration	Float	Total time duration of a segment in seconds.
GPS Coordinates	Dummy	Latitude and longitude coordinates
Time Stamp	Dummy	Time information for each GPS point
Mode Label	Dummy	Ground truth label for transportation mode (e.g., Walk, Run, Bike, Car, Bus)

$$TR_i = \frac{b_{i+1} - b_i}{\Delta t_i} \quad 1 \leq i \leq N; TR_1 = TR_N = 0 \tag{9}$$

Where  $A_p$  and  $AvgS_p$  represent an acceleration of the point and the average speed of the points in each segment,  $n$  means the number of positional recordings in each segment, and the Jerk (the rate of change in acceleration) and turning rate values, respectively.

The bearing rate is calculated according to Eq. 10 to 12, by using two sequential positional recordings information:

$$y = \text{sine} [P_2(\text{long}) - P_1(\text{long})] * \text{cosine} [P_2(\text{lat})] \tag{10}$$

$$x = \text{cosine} [P_1(\text{lat})] * [P_2(\text{lat})] - \text{sine} [P_1(\text{lat})] * \text{cosine} [P_2(\text{lat})] * \text{cosine} [P_2(\text{long}) - P_1(\text{long})] \tag{11}$$

$$\text{Bearing}_{(P_1)} = \text{arctan}(y, x) \tag{12}$$

Another essential feature is a segment duration that is calculated based on the time difference of the first and the endpoint of the segment. Moreover, the distance between two sequential positional recordings and the total distance of each segment is calculated according to the following equations:

$$D_{(p_1,p_2)} = \text{cosine} [\text{sine} [P_1 (\text{lat})] * \text{sine} [P_2 (\text{lat})] + \text{cosine} [P_2 (\text{lat})] * \text{cosine} [P_2 (\text{long}) - P_1 (\text{long})]] * \frac{180}{\pi} * 60 \quad (13)$$

$$D_T = \sum_{i=1}^n D_i \quad (14)$$

In addition to the extracted and computed features from the raw GPS tracking dataset, we incorporate GIS information layers including road and rail networks, bus and train stations to enable enhanced dataset visualization. Specifically, these GIS layers provide a more comprehensive perspective on train and bus segments. After obtaining the clean and validated GPS dataset, the dataset was used for feeding into the LSTM Autoencoder, which is described in detail in the next subsection.

### Semi-supervised learning with stacking LSTM

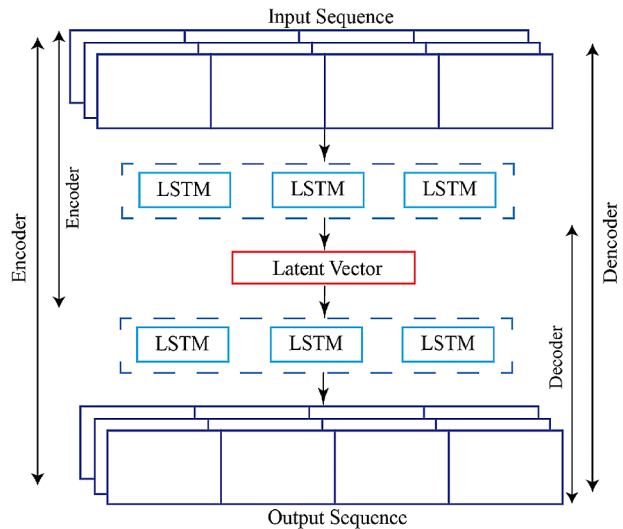
In most studies, the conventional approach involves the extraction or computation of various features, which are then directly utilized for transportation mode detection. This method proves effective when trajectory data are entirely labelled with corresponding transportation mode information, facilitating straightforward network training. The input data for the model is reshaped to incorporate time step information. This restructuring is crucial for extracting temporal correlations from the data, a vital aspect in travel mode identification. By mapping the input data into an R-dimensional time series using LSTM, the model effectively captures these temporal dynamics, which are essential for improving the accuracy of the model in identifying travel modes. However, it often remains unclear which of these extracted features holds the utmost effectiveness in discerning transportation modes. Nevertheless, real-world scenarios frequently involve a majority of unlabelled GPS data, rendering the use of a fully labelled dataset impractical.

To address this challenge, a deep learning algorithm can be employed to amalgamate and distil pertinent information, ultimately reaching a conclusive decision. In this context, we employ a straightforward yet robust solution: a simple LSTM autoencoder. This approach allows us to extract latent information from the computed features. While not explicitly performing feature selection, this architecture implicitly learns feature representations, emphasizing those that are most salient for the task at hand. Post-training feature importance or relevance analysis can further substantiate our conclusions regarding feature effectiveness.

### LSTM autoencoder

LSTM was proposed by Hochreiter and Schmidhuber (1997), and is one of the most commonly used data mining methods with a modern architecture of neural networks. While standard neural networks attempt temporal correlation within the training data, the LSTM algorithm learns from such correlation to extract latent information, due to the special design of states that propagate over time. The LSTM autoencoder is an implementation of an autoencoder that approves time series sequential data and uses LSTM architecture in its encoder and decoder. The encoder LSTM reads the input sequence, and the model stores the latent information as a vector, which represents the entire original dataset. Figure 3 represents the architecture of an LSTM autoencoder and the data flow within it.

**Fig. 3** The architecture of a simple LSTM autoencoder



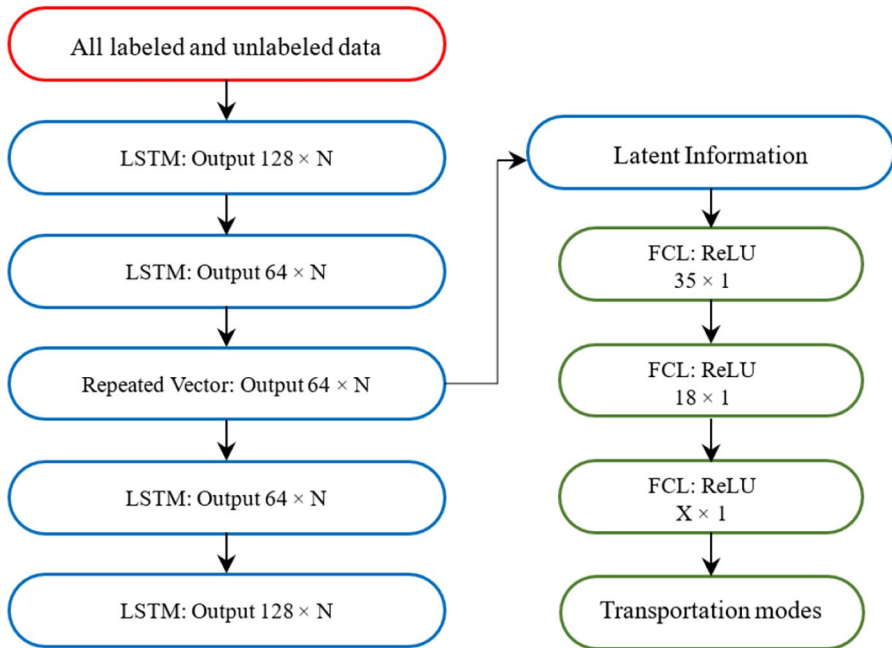
### Single classifier

In order to regenerate the input data, the vector of hidden information can be passed as the input to the decoder part of the LSTM autoencoder. After extracting the latent information, it needs to build a fully connected neural network that uses the inputs and detects the transportation modes. The fully connected neural network can be trained by using categorical cross-entropy as a loss function, since the purpose of the model is multiclass classification. Moreover, in order to optimize the model, an Adam optimizer is used. Figure 4 presents the architecture of the model. Figure 4 offers an illustrative overview of our model's structural components, showcasing how latent information is leveraged and transformed to make accurate predictions regarding transportation modes.

The stack network exploits the latent information with different hyperparameter configurations. Three fully connected layers are employed, which take latent information as input for the network. 35 and 18 neurons are employed in the first two ReLU activated layers. In the last layer with  $X$  number of neurons, the softmax activation function is employed to identify the travel modes, where  $X$  is the number of transportation modes.

### Results

For analyzing the data, the Python 3.7 programming language was used. The LSTM autoencoder architectures were implemented in Keras, a deep learning library in Python, utilizing the TensorFlow backend with the support of only CPU. 80% of the whole labeled data was randomly selected as training dataset while the rest was used as testing data. The performance evaluation of an LSTM autoencoder model needs to be done only on the test dataset that does not affect the training process.



**Fig. 4** The architecture of the proposed model for detecting transportation modes

## Dataset processing

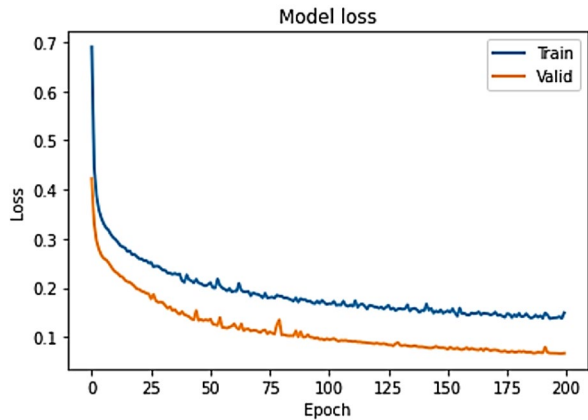
In Sect. 4.1, we provided insights into our data collection process. The raw GPS tracking data was collected from a subset of our volunteer group, which consisted of a total of 91 individuals, over a span of one year. During this data collection phase, 15 participants contributed to the labeling process. We employed a user-based approach, where participants were asked to manually record the mode of transportation for each segment of their trajectory. In this dataset, each trajectory was represented as a sequence of GPS positional recordings. In total, we gathered 8,226 segments that were labeled by these 15 participants. These labels indicated the transportation mode used for each corresponding trajectory segment. Specifically, participants assigned one of five following transportation modes: walk, bike, bus, car, or train to their recorded segments. To further process and analyze the raw GPS data, we applied the pruned exact linear time (PELT) algorithm to detect change points in trips based on speed and bearing rates (Killick et al. 2012). The PELT algorithm is used for changepoint detection in GPS tracking data. It efficiently identifies points where there's a significant change in speed and bearing rates, helping segment trajectories into different modes of transportation (walk, bike, bus, car, or train). It has a linear time complexity, making it computationally efficient for large datasets.

This resulted in a total of 612,088 segmented data points. After segmenting the data, we calculated all features listed in Table 2 for each positional recording as was described in Sect. 3.1. Table 3 below provides an overview of the total number of labeled segments

**Table 3** Number of samples, maximum speed, and acceleration associated with each transportation mode

Transportation mode	Number of points	Number of segments	Maximum speed (m/s)	Maximum acceleration (m/s <sup>2</sup> )
Walk	15,269	2,839	19	1.6
Bike	10,850	1,351	49	5.8
Bus	11,624	874	110	5
Car	12,636	1,491	129	7.8
Train	15,736	1,671	136	7.2

**Fig. 5** Changes in loss function over numbers of iteration



**Table 4** Performance comparison of with varying amounts of labeled data

Portion (%)	Fold 1 (%)	Fold 2 (%)	Fold 3 (%)	Fold 4 (%)	Fold 5 (%)	Accuracy (%)
10	89.04	89.78	89.56	89.78	89.18	89.47
20	91.87	91.98	91.76	89.94	91.15	91.34
50	93.38	92.30	92.15	93.33	93.60	92.95
100	94.06	94.06	93.81	93.72	94.03	93.94

and point data for each type of transportation mode, offering insight into the distribution of ground truth data used for our analysis.

**LSTM autoencoder performance**

The LSTM Autoencoder with the architecture shown in Fig. 3 is trained with all labeled and unlabeled point data. Figure 5 shows the changes in the loss over 200 iterations, and the LSTM autoencoder algorithm predicts the input values with less than 0.1 error. Therefore, it can be concluded that the obtained latent information can be reliable and can function as a good representative of the dataset as the input values.

A single deep neural network with three hidden layers and two input and output layers is employed to identify the transportation modes. In order to avoid biased results, K-fold cross-validation with five numbers of splits and only one iteration is applied. Table 4 shows the model performance with a different portion of labeled data.

Table 4 shows the overall model performance using different portions of labeled data. As can be seen, the performance of the model with only using 10% of labeled data is over 89%. By increasing the portion labeled, the performance of the model increased significantly. The achieved results reveal that the proposed model can identify the correct transportation modes with a high level of accuracy. Moreover, the results proven that the proposed model could work well with over-fitting issues prominently.

Table 5 shows the model accuracy concerning using different proportions of the labeled dataset. The table shows that with using only 10% of the labeled dataset, the model accuracy is 89.47%, while by increasing the use of the labeled dataset (20%), the accuracy values reach 91.34%. Moreover, 50% and 100% of partition of a labeled dataset are used, and the accuracy values reach 92.95% and 93.94%, respectively. The result indicates that by only using 20% of the labeled dataset, the proposed model can achieve over 91% accuracy in detecting the correct transportation modes. Moreover, the results confirmed that the LSTM Autoencoder does not prominently suffer from over-fitting issues. The loss and accuracy of training the deep neural network on a totally full labeled dataset over 100 iterations for each fold were examined. The loss value at the end of the 100 iterations per fold was below 0.2, which showed how the model overperformed in predicting the five transportation modes. Table 5 shows how well the proposed algorithm performs in identifying a specific transportation mode by using different portions of the labeled dataset.

The table shows when the training dataset is 10% of the labeled dataset, the performance of the proposed model is significant, almost 90%, in detecting the correct modes. The model performed well in detecting walk, bike, and train modes with a precision level of above 90%. However, the model performance in detecting car and bus modes using a 10% labeled dataset, compared to other modes, is also acceptable (over 80%). By increasing the amount of labeled data (20%), the performance of the proposed model is increased to 91.34%. The precision value for walk mode was approximately the same; however, for other modes, the values improved significantly.

**Table 5** Confusion matrix of the proposed model with using 10% and 20% of labeled data

10% of labeled data		Predicted mode						Sum	Recall (%)
		Bike	Bus	Car	Train	Walk			
Actual mode	Bike	984	36	28	11	0	1,059	92.92	
	Bus	67	969	137	22	0	1,195	81.09	
	Car	26	171	1,026	68	1	1,292	79.41	
	Train	24	22	81	1,427	0	1,554	91.83	
	Walk	0	0	2	0	1,509	1,511	99.87	
	Sum	1,101	1,198	1,274	1,528	1,510	6,611	-	
	Precision (%)	89.37	80.88	80.53	93.39	99.93	-	-	
20% of labeled data		Predicted mode					Sum	Recall (%)	
		Bike	Bus	Car	Train	Walk			
Actual mode	Bike	2,075	66	33	24	0	2,198	94.40	
	Bus	122	1,860	239	20	0	2,241	83.00	
	Car	31	305	2,168	84	2	2,590	83.71	
	Train	21	32	164	2,896	0	3,113	93.03	
	Walk	0	0	2	0	3,079	3,081	99.94	
	Sum	2,249	2,263	2,606	3,024	3,081	13,223	-	
	Precision (%)	92.26	82.19	83.19	95.77	99.94	-	-	

Table 6 shows the confusion matrix of the proposed model using 50% and 100% of labeled data. As can be seen, the precision and recall values with using 100% of labeled data are increased, compared to using 50% labeled data. The precision values are improved by around 2% for bus and car modes. Nevertheless, the obtained precision for walk, bike, and train are not significant in comparison to using only 50% labeled data because it performs as well if computing time is considered.

The overall results show that the walk mode is detected more accurately by the model since it has a larger amount of point data in the training dataset. However, bus and car modes are easily misclassified by the model since these two modes have the same behavior in the motion, speed, and road network. Moreover, the model accuracy in detecting the train mode compared to the other motorized modes was higher due to more labeled data than bus, and car modes. The bus mode is also classified as a bike mode, due to the low speed of the bus in city traffic, and the acceleration and turning rate when the bus is close to traffic lights. There is also misclassification between car and train modes because of some similarities in acceleration and deceleration when the trains arrive at train stations. Moreover, the rail network in the study area has some similarities with the road network.

### Comparison with other studies

To demonstrate the superiority of our proposed model, we conducted a comparative analysis of its performance against other studies that have utilized deep learning models with GPS tracking datasets for transportation mode detection.

In this section, we focus on a series of baseline studies and compare their accuracy level of detection on different types of modes, including semi-supervised convolutional Auto-encoder (SECA)(Dabiri et al. 2019), Semi-two-steps (Dabiri et al. 2019), Semi-pseudo-label(Dabiri et al. 2019), Generative Adversarial Network (GAN) (Yazdizadeh et al. 2021), LSTM (Asci and Guvensan 2019), and semi-supervised deep ensemble learning algorithm (Yu 2020). Additionally, some traditional machine learning algorithms studies are also used

**Table 6** Confusion matrix of the proposed model with using 50% and 100% of labeled data

50% of labeled data		Predicted mode						
		Bike	Bus	Car	Train	Walk	Sum	Recall (%)
Actual mode	Bike	5,201	122	68	47	0	5,438	95.64
	Bus	211	5,074	535	45	0	5,865	86.51
	Car	62	653	5,413	164	1	6,293	86.02
	Train	34	67	315	7,461	1	7,878	94.71
	Walk	0	0	5	0	7,578	7,583	99.93
	Sum	5,508	5,916	6,336	7,717	7,580	33,057	-
	Precision (%)	94.43	85.77	85.43	96.68	99.97	-	-
100% of labeled data		Predicted mode						
		Bike	Bus	Car	Train	Walk	Sum	Recall (%)
Actual mode	Bike	10,508	158	132	52	0	10,850	96.85
	Bus	368	10,444	764	48	0	11,624	89.85
	Car	159	1,260	10,858	356	3	12,636	85.93
	Train	53	93	552	15,035	3	15,736	95.55
	Walk	0	0	7	0	15,262	15,269	99.95
	Sum	11,088	11,955	12,313	15,491	15,268	66,115	-
	Precision (%)	94.77	87.36	88.18	97.06	99.96	-	-

in the comparison, including decision tree algorithms, K- nearest neighbor, super vector machine, random trees, multi-layer perceptron (Dabiri and Heaslip 2018). The results of the performance of the models are shown and summarized in Table 7. In this table, the value is developed by our implementations shown as “proposed,” and the performance of other methods are also presented with reference.

Based on the table, it is evident that the proposed model in this study consistently outperforms other semi-supervised transportation mode detection studies across different labelled data percentages (10%, 20%, 50%, and 100%). Notably, even with only 10% labelled data, our model achieves an accuracy rate of 89.5%, highlighting its robustness and efficiency. Comparing our model to the second-best performing study, Deep ensemble learning by Yu (2020), reveals a significant difference in complexity. While Yu (2020) employed six stacked LSTM layers and four different deep neural networks, our approach utilizes a simpler LSTM autoencoder with four layers (two for encoding and two for decoding) in conjunction with a deep neural network featuring three hidden layers. Despite its reduced complexity, our method achieves superior accuracy results. This underscores the effectiveness of our approach in achieving high accuracy with limited labelled data, making a fully labelled dataset unnecessary for achieving exceptional performance.

While all algorithms demonstrate similar performance in transportation mode detection, there are several reasons to consider our approach. Firstly, our algorithm is designed to be computationally efficient, making it suitable for real-time or resource-constrained applications. In contrast, deep ensemble learning often involves complex models with higher computational costs. Secondly, our model achieves comparable accuracy while using a smaller amount of labelled data. This is especially valuable when labelled data is scarce or expensive to acquire. Additionally, our algorithm’s architecture is more streamlined and may require less fine-tuning compared to the ensemble approach. Lastly, our model’s interpretability and feature extraction capabilities can provide valuable insights into which features are most influential in mode detection. These considerations make our algorithm a compelling choice, especially when balancing performance, efficiency, and interpretability.

**Table 7** Performance comparison among transportation mode detection models under varying percentages of labeled data

Percent of labeled data	10% labeled data	20% labeled data	50% labeled data	100% labeled data
Models	Accuracy (%)			
Proposed model	89.5	91.3	92.9	93.9
SECA (Dabiri et al. 2019)	72.9	74.1	77.2	78.2
Semi-two-step (Dabiri et al. 2019)	64.1	66.2	68.8	70.4
Semi-pseudo-label (Dabiri et al. 2019)	68.9	70.3	71.2	75.4
GAN (Yazdizadeh et al. 2021)	78.4	79.9	81.9	82.6
KNN (Dabiri et al. 2019)	61.1	61.8	62.9	63.4
SVM (Dabiri et al. 2019)	61.7	62.1	63.1	63.5
DT (Dabiri et al. 2019)	77.1	78.2	78.8	79.4
MLP (Dabiri et al. 2019)	48.4	49.2	52.1	54.9
CNN (Dabiri et al. 2019)	86.8	87.2	87.9	88.1
RF (Yu 2020)	83.9	84.3	85.8	87.7
LSTM (Asci and Guvensan 2019)	78.9	79.1	80.5	81.2
Deep ensemble learning (Yu 2020)	88.7	89.4	90.1	91.2



## Conclusions

In this paper, a novel transportation mode detection model based on semi-supervised deep learning is proposed. The model is based on a combination of a long short-term memory (LSTM) Autoencoder and a deep neural network algorithm. The deep neural network can extract the latent information from the applied LSTM autoencoder to generate the label for the unlabeled dataset, based on the existing information of GPS labeled data. This combination does not require fully labelling of the dataset and can show which extracted features are most effective in detecting transportation modes. This approach can use the unlabeled human mobility GPS trajectories to improve system performance. The proposed model is applied to a large volume of GPS tracking dataset for detecting five transportation modes. The results show that it can detect the correct transportation modes effectively and with a high accuracy of 93.94%. However, note that the proposed model in this paper used only 6% of labeled data. This approach offers several advantages, including the ability to make accurate predictions without the need for fully labeled datasets and the identification of key features for transportation mode detection. To assess the performance of the proposed model, two aspects were considered. First, the accuracy of the model was evaluated by using different portions of the labeled data. Second, a comparison analysis with other studies applied semi-supervised learning was conducted. The result showed that the proposed model performed well in predicting the correct transportation modes for unlabeled data and outperformed other semi-supervised learning models.

The high accuracy achieved in our study can indeed be attributed, in part, to the homogeneous nature of our dataset, where each participant contributed three weeks of data. This extended data collection period per person allowed our algorithm to learn individual transportation behaviours thoroughly, thereby enhancing its accuracy in mode detection. We acknowledge that in real-world scenarios, data collection periods may vary widely, which can impact the model accuracy. However, our approach is designed to be able to adapt and fit for diverse data volume and less homogeneous datasets, which makes it a versatile solution for various applications. While the homogeneity of our sample played a role in achieving high accuracy, the ability that the model can adapt and can be generalized under different scenarios remains a notable strength.

The proposed algorithm is not solely reliant on labelled data; it is also built to work with unlabelled trajectory data from sources like smartphones and drones. Its versatility means it provides better accuracy when labels are available but remains useful with unlabelled data through its semi-supervised approach. While our dataset, consisting of 91 participants, may have limitations in broader applications, we believe our approach can adapt to various real-world scenarios, making it practical beyond the scope of our dataset.

However, it's important to acknowledge certain limitations associated with our research and the data used. First, our data primarily comes from GPS tracking data collected in a specific location (Borlänge, Sweden) over a limited timeframe. While our model has demonstrated high accuracy, further investigation is needed to determine how well it can be generalized under diverse geographic regions and different time periods. Additionally, the accuracy of our model depends on the initial labels provided by the 10 participants. Human labeling can introduce errors, and the impact of labeling inaccuracies on our model's performance should be considered.

The practical implications of our analysis are still noteworthy. Firstly, our model's ability to achieve accurate results with minimal labeled data has significant implications for reducing the cost and time associated with data labeling in transportation mode detection applications. This can streamline the development of transportation-related systems and services. Secondly, by effectively utilizing unlabeled GPS data, our model enhances system performance in applications such as urban planning, traffic management, and environmental monitoring. Accurate transportation mode detection is crucial for making informed decisions in these domains. Lastly, while our research was conducted in a specific context, the model's architecture and approach can be applied to various locations and datasets, potentially benefiting transportation-related studies worldwide.

In conclusion, our research not only offers a promising approach to transportation mode detection but also highlights the need for further investigation into the model's generalization potential, labeling accuracy, and applicability to diverse datasets. The practical implications of our work also extend to more efficient data labeling and improved system performance in transportation-related applications, showcasing the practical relevance of our findings.

**Author contributions** Paria Sadeghian: Conceptualization, Methodology, Data curation, Writing- Original draft preparation. Arman Golshan: Methodology, Data curation, Visualization. Xiaoyun Zhao: Conceptualization, Writing- Reviewing and Editing. Johan Håkansson: Conceptualization, Writing- Reviewing and Editing

**Funding** Open access funding provided by Dalarna University.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alam, M.G.R., Haque, M., Hassan, M.R., Huda, S., Hassan, M.M., Strickland, F.L., AlQahtani, S.A.: Feature cloning and feature Fusion Based Transportation Mode Detection using convolutional neural network. *IEEE Trans. Intell. Transp. Syst.* **24**, 4671–4681 (2023)
- Asci, G., Guvensan, M.A.: A novel input set for lstm-based transport mode detection. In: 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). pp. 107–112. IEEE (2019)
- Bachir, D., Khodabandelou, G., Gauthier, V., Yacoubi, E., Puchinger, M.: Inferring dynamic origin-destination flows by transport mode using mobile phone data. *Transp. Res. Part. C Emerg. Technol.* **101**, 254–275 (2019)
- Dabiri, S., Heaslip, K.: Inferring transportation modes from GPS trajectories using a convolutional neural network. *Transp. Res. part. C Emerg. Technol.* **86**, 360–371 (2018). <https://doi.org/10.1016/j.trc.2017.11.021>
- Dabiri, S., Lu, C., Heaslip, K., Reddy, C.K., Member, S.: Semi-Supervised Deep Learning Approach for Transportation Mode Identification Using GPS Trajectory Data. X (2019). <https://doi.org/10.1109/TKDE.2019.2896985>

- Dabiri, S., Marković, N., Heaslip, K., Reddy, C.K.: A deep convolutional neural network based approach for vehicle classification using large-scale GPS trajectory data. *Transp. Res. Part. C Emerg. Technol.* **116**, 102644 (2020). <https://doi.org/10.1016/j.trc.2020.102644>
- Dutta, S., Patra, B.K.: Inferencing transportation mode using unsupervised deep learning approach exploiting GPS point-level characteristics. *Appl. Intell.* **53**, 12489–12503 (2023)
- Feng, T., Timmermans, H.J.P.: Comparison of advanced imputation algorithms for detection of transportation mode and activity episode using GPS data. **1060**, (2016). <https://doi.org/10.1080/03081060.2015.1127540>
- Frappart, F., Bourrel, L.: *The Use of Remote Sensing in Hydrology*. MDPI-Multidisciplinary Digital Publishing Institute (2018)
- Giri, S., Brondeel, R., Aarbaoui, E., Chaix, T.: Application of machine learning to predict transport modes from GPS, accelerometer, and heart rate data. *Int. J. Health Geogr.* **21**, 19 (2022)
- Gong, H., Chen, C., Bialostozky, E., Lawson, C.T.: Computers, Environment and Urban Systems a GPS / GIS method for travel mode detection in New York City. *Comput. Environ. Urban Syst.* **36**, 131–139 (2012). <https://doi.org/10.1016/j.compenvurbysys.2011.05.003>
- Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997)
- Jackson, B., Scargle, J.D., Barnes, D., Arabhi, S., Alt, A., Gioumouisis, P., Gwin, E., Sangtrakulcharoen, P., Tan, L., Tsai, T.T.: An algorithm for optimal partitioning of data on an interval. *IEEE Signal. Process. Lett.* **12**, 105–108 (2005)
- Killick, R., Fearnhead, P., Eckley, I.A.: Optimal detection of changepoints with a linear computational cost. *J. Am. Stat. Assoc.* **107**, 1590–1598 (2012)
- Kumar, S., Damaraju, A., Kumar, A., Kumari, S., Chen, C.-M.: LSTM network for transportation mode detection. *J. Internet Technol.* **22**, 891–902 (2021)
- Lari, Z.A., Golroo, A.: Automated transportation mode detection using smart phone applications via machine learning: Case study mega city of Tehran. In: *Proceedings of the Transportation Research Board 94th Annual Meeting*, Washington, DC, USA. pp. 11–15 (2015)
- Lee, K., Kwan, M.P.: Automatic physical activity and in-vehicle status classification based on GPS and accelerometer data: A hierarchical classification approach using machine learning techniques. *Trans. GIS.* **22**, 1522–1549 (2018). <https://doi.org/10.1111/tgis.12485>
- Li, J., Pei, X., Wang, X., Yao, D., Zhang, Y., Yue, Y.: Transportation mode identification with GPS trajectory data and GIS information. *Tsinghua Sci. Technol.* **26**, 403–416 (2021). (a)
- Li, R., Yang, Z., Pei, X., Yue, Y., Jia, S., Han, C., He, Z.: A novel one-stage approach for pointwise transportation mode identification inspired by point cloud processing. *Transp. Res. Part. C Emerg. Technol.* **152**, 104127 (2023)
- Li, Z., Xiong, G., Wei, Z., Lv, Y., Anwar, N., Wang, F.-Y.: A semisupervised end-to-end framework for transportation mode detection by using gps-enabled sensing devices. *IEEE Internet Things J.* **9**, 7842–7852 (2021)(b)
- Lin, M., Hsu, W., Lee, Z.Q.: Detecting modes of transport from unlabelled positioning sensor data. *J. Locat Based Serv.* **7**, 272–290 (2013). <https://doi.org/10.1080/17489725.2013.819128>
- Ma, Y., Guan, X., Cao, J., Wu, H.: A multi-stage fusion network for transportation mode identification with varied scale representation of GPS trajectories. *Transp. Res. Part. C Emerg. Technol.* **150**, 104088 (2023)
- Markos, C., Yu, J.J.Q.: Unsupervised Deep Learning for GPS-Based Transportation Mode Identification. 1–6 (2020). <https://doi.org/10.1109/itsc45102.2020.9294673>
- Marra, A.D., Sun, L., Corman, F.: The impact of COVID-19 pandemic on public transport usage and route choice: Evidences from a long-term tracking study in urban area. *Transp. Policy.* **116**, 258–268 (2022)
- Moreau, H., Vassilev, A., Chen, L.: The devil is in the details: An efficient convolutional neural network for transport mode detection. *IEEE Trans. Intell. Transp. Syst.* **23**, 12202–12212 (2021)
- Namdarpour, F., Mesbah, M., Gandomi, A.H., Assemi, B.: Using genetic programming on GPS trajectories for travel mode detection. *IET Intell. Transp. Syst.* **16**, 99–113 (2022)
- Nick, T., Coersmeier, E., Geldmacher, J., Goetze, J.: Classifying means of transportation using mobile sensor data. In: *The 2010 International Joint Conference on Neural Networks (IJCNN)*. pp. 1–6. IEEE (2010)
- Nitsche, P., Widhalm, P., Breuss, S., Brändle, N., Maurer, P.: Supporting large-scale travel surveys with smartphones – A practical approach. *Transp. Res. Part. C.* **43**, 212–221 (2014). <https://doi.org/10.1016/j.trc.2013.11.005>
- Patterson, D.J., Liao, L., Fox, D., Kautz, H.: Inferring High-Level Behavior from low-level sensors. In: Dey, A.K., Schmidt, A., McCarthy, J.F. (eds.) *UbiComp 2003: Ubiquitous Computing*, pp. 73–89. Springer Berlin Heidelberg, Berlin, Heidelberg (2003)
- Raza, S., Al-Kaisy, A., Teixeira, R., Meyer, B.: The role of GNSS-RTN in Transportation Applications. *Encyclopedia.* **2**, 83 (2022)

- Reddy, S., Mun, M.I.N., Burke, J., Estrin, D., Hansen, M., Srivastava, M.: Using mobile phones to determine transportation modes. *ACM Trans. Sens. Networks*. **6**, 1–27 (2010). <https://doi.org/10.1145/1689239.1689243>
- Rezaie, M., Patterson, Z., Yazdizadeh, A.: Semi-supervised Travel Mode Detection from Smartphone Data. (2017). <https://doi.org/10.1109/ISC2.2017.8090800>
- Roy, A., Fuller, D., Nelson, T., Kedron, P.: Assessing the role of geographic context in transportation mode detection from GPS data. *J. Transp. Geogr.* **100**, 103330 (2022)
- Sadeghian, P., Håkansson, J., Zhao, X.: Review and evaluation of methods in transport mode detection based on GPS tracking data. *J. Traffic Transp. Eng. (English Ed.)* **8**, 467–482 (2021). <https://doi.org/10.1016/j.jtte.2021.04.004>
- Sadeghian, P., Zhao, X., Golshan, A., Håkansson, J.: A stepwise methodology for transport mode detection in GPS tracking data. *Travel Behav. Soc.* **26**, 159–167 (2022). <https://doi.org/10.1016/j.tbs.2021.10.004>
- Shen, L., Stopher, P.R.: Review of GPS Travel Survey and GPS Data- Processing methods. *Transp. Rev.* **0**, 1–19 (2014). <https://doi.org/10.1080/01441647.2014.903530>
- Tamim Kashifi, M., Jamal, A., Samim Kashefi, M., Almoshaogeh, M., Masiur Rahman, S.: Predicting the travel mode choice with interpretable machine learning techniques: A comparative study. *Travel Behav. Soc.* **29**, 279–296 (2022). <https://doi.org/10.1016/j.tbs.2022.07.003>
- Vinayaraj, P., Mede, K.: Multi-branch deep learning based transport mode detection using weakly supervised labels. *Int. Arch. Photogramm Remote Sens. Spat. Inf. Sci.* **48**, 525–530 (2022)
- Vincenty, T.: Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. *Survey Rev.* **23**(176), 88–93 (1975)
- Weinstein, P.A.G.J.S., Labrador, S.J.B.M.A., Perez, P.L.W.N.L.G.R.: Automating mode detection for travel behaviour analysis by using global positioning systems- enabled mobile phones and neural networks. (2010). <https://doi.org/10.1049/iet-its.2009.0029>
- Wu, L., Yang, B., Jing, P.: Travel mode detection based on GPS raw data collected by smartphones: A systematic review of the existing methodologies. *Information*. **7**, 67 (2016). <https://doi.org/10.3390/info7040067>
- Xiao, Z., Wang, Y., Fu, K., Wu, F.: Identifying different transportation modes from trajectory data using tree-based ensemble classifiers. *ISPRS Int. J. Geo-Information*. **6**, 57 (2017)
- Xu, Y., Zou, D., Park, S., Li, Q., Zhou, S., Li, X.: Understanding the movement predictability of international travelers using a nationwide mobile phone dataset collected in South Korea. *Comput. Environ. Urban Syst.* **92**, 101753 (2022). <https://doi.org/10.1016/j.compenvurbsys.2021.101753>
- Yao, Y., Zhang, H., Chen, Q.: Trip segmentation and mode detection for human mobility data. In: *Handbook of Mobility Data Mining*, pp. 97–115. Elsevier (2023)
- Yazdizadeh, A., Patterson, Z., Farooq, B.: Ensemble convolutional neural networks for mode inference in smartphone travel survey. *IEEE Trans. Intell. Transp. Syst.* **21**, 2232–2239 (2019). <https://doi.org/10.1109/TITS.2019.2918923>
- Yazdizadeh, A., Patterson, Z., Farooq, B.: Semi-supervised gans to infer travel modes in GPS trajectories. *J. Big Data Anal. Transp.* 1–11 (2021)
- Yu, J.J.Q.: Semi-supervised deep ensemble learning for travel mode identification. *Transp. Res. Part. C*. **112**, 120–135 (2020). <https://doi.org/10.1016/j.trc.2020.01.003>
- Zhagyparova, K., Bader, A., Kouzayha, N., ElSawy, H., Al-Naffouri, T.: Transportation Mode Recognition based on Cellular Network Data. In: 2023 IEEE International Conference on Smart Mobility (SM). pp. 132–137. IEEE (2023)
- Zhang, X., Gao, Y., Wang, X., Feng, J., Shi, Y.: GeoSDVA: A semi-supervised Dirichlet Variational Auto-encoder Model for Transportation Mode Identification. *ISPRS Int. J. Geo-Information*. **11**, 290 (2022)
- Zhao, X., Carling, K., Håkansson, J.: An evaluation of the reliability of GPS-based transportation data. In: *IAC (International Academic Conference on Transport, Logistics, Tourism and Sport Science) in Vienna November 24–25 2017* pp. 323–334 (2017)
- Zheng, Y., Li, Q., Chen, Y., Xie, X., Ma, W.-Y.: Understanding mobility based on GPS data. In: *Proceedings of the 10th international conference on Ubiquitous computing*. pp. 312–321 (2008)(a)
- Zheng, Y., Liu, L., Wang, L., Xie, X.: Learning transportation mode from raw gps data for geographic applications on the web. In: *Proceedings of the 17th international conference on World Wide Web*. pp. 247–256 (2008)(b)

**Paria Sadeghian** is a PhD candidate at Dalarna University, Sweden. She received her M.S. degree in industrial engineering from Khatam University. Her research interests are in transport mode detection, urban mobility, travel behaviour, and data analytics.

**Arman Golshan** is a master's student at Dalarna University, Sweden. He received his bachelor's degree in information technology from Qazvin Azad University. His research interests are in transportation, urban mobility, and data analytics.

**Mia Xiaoyun Zhao** is an Associate Professor in microdata analysis at Dalarna University and an affiliated researcher at the Integrated Transport Research Lab at KTH Royal Institute of Technology, Sweden. Her research focuses on sustainable transport solutions, system-level impacts of smart mobility, urban transport planning, transport simulation, and data analytics.

**Johan Håkansson** is a full professor in microdata analysis at Dalarna University, Sweden. His research interests are in urban and regional mobility, last-mile logistics, location analysis, and data analytics.