Transportation Research Part C 86 (2018) 360–371

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Inferring transportation modes from GPS trajectories using a convolutional neural network

Sina Dabiri*, Kevin Heaslip

Charles E. Via, Jr. Department of Civil and Environmental Engineering, 301, Patton Hall, Virginia Tech, Blacksburg, VA 24061, United States

ARTICLE INFO

Keywords:
Convolutional neural network
Deep learning
GPS data
Transportation mode Inference

ABSTRACT

Identifying the distribution of users' transportation modes is an essential part of travel demand analysis and transportation planning. With the advent of ubiquitous GPS-enabled devices (e.g., a smartphone), a cost-effective approach for inferring commuters' mobility mode(s) is to leverage their GPS trajectories. A majority of studies have proposed mode inference models based on hand-crafted features and traditional machine learning algorithms. However, manual features engender some major drawbacks including vulnerability to traffic and environmental conditions as well as possessing human's bias in creating efficient features. One way to overcome these issues is by utilizing Convolutional Neural Network (CNN) schemes that are capable of automatically driving high-level features from the raw input. Accordingly, in this paper, we take advantage of CNN architectures so as to predict travel modes based on only raw GPS trajectories, where the modes are labeled as walk, bike, bus, driving, and train. Our key contribution is designing the layout of the CNN's input layer in such a way that not only is adaptable with the CNN schemes but represents fundamental motion characteristics of a moving object including speed, acceleration, jerk, and bearing rate. Furthermore, we ameliorate the quality of GPS logs through several data preprocessing steps. Using the clean input layer, a variety of CNN configurations are evaluated to achieve the best CNN architecture. The highest accuracy of 84.8% has been achieved through the ensemble of the best CNN configuration. In this research, we contrast our methodology with traditional machine learning algorithms as well as the seminal and most related studies to demonstrate the superiority of our framework.

1. Introduction

Travel mode choice is one of the principal traveler's behavior attributes in travel demand analysis, transport planning, and traffic management. By inferring the travel mode distribution, transportation agencies are able to generate appropriate strategies to alleviate users' travel time, traffic congestion, and air pollution. For example, a clear benefit of travel mode analysis is to identify regions with high auto dependency and encourage public transport ridership by improving transit systems (Eluru et al., 2012). Furthermore, suitable policies such as HOV lanes can be taken during the peak-period congestion according to existing mode shares. The knowledge of the transport mode choice has traditionally been obtained through household surveys or phone interviews. However, interviewing households is a time-consuming and expensive method that usually results in a low response rate and incomplete information, which calls for a cost-effective technology such as Global Positioning System (GPS) that is able to collect travel data while reducing labor and time costs.

GPS is a ubiquitous positioning tool that records spatiotemporal information of moving objects carrying a GPS-enabled device

* Corresponding author.
E-mail addresses: sina@vt.edu (S. Dabiri), kheaslip@vt.edu (K. Heaslip).

https://doi.org/10.1016/j.trc.2017.11.021
Received 14 August 2017; Received in revised form 26 November 2017; Accepted 26 November 2017
0968-090X/ Published by Elsevier Ltd.

(e.g., a smartphone). The main advantageous of smart phones, compared to other GPS-equipped devices, is its enormous market penetration rate in a large number of countries and being relatively close to users nearly all of the time. As a consequence, such a dominant and area-wide sensing technology is capable of creating massive trajectory data of vehicles and people. A GPS trajectory, also called movement, of an object is constructed by connecting GPS points of their GPS-enabled device. A GPS point, here, is denoted as (*lat*, *long*, *t*), where *lat*, *long*, and *t* are latitude, longitude, and timestamp, respectively. The study of individuals' mobility patterns from GPS datasets has led to a variety of behavioral applications including learning significant locations, anomaly detection, location-based activity recognition, and identification of transport modes (Lin and Hsu, 2014), in which the latter is the focus of this study. Nonetheless, GPS devices can only record time and positional characteristics of travels without any explicit information on utilized transport modes. This necessitates employing data processing and mining algorithms to extract hidden knowledge about transport modes from raw GPS data.

Many of proposed inference models on detecting transport modes by means of only GPS sensors include two steps. In the first step, a pool of attributes (e.g., velocity, acceleration, heading change rate, and stop rate) (Zheng et al., 2008a,b), are computed from GPS logs. In the second step, the extracted features are fed into a learning algorithm to estimate the transportation mode. Unlike many classification problems that a majority of features have already been computed and included in the dataset, raw GPS trajectories contain only a series of chronologically ordered points without any explicit features such as speed and acceleration. This fact has required researchers to manually identify and formulate a set of features before using a machine learning technique for the classification task. However, the hand-crafted features may not necessarily distinguish between various transportation modes since they are vulnerable to traffic and environmental conditions (Zheng et al., 2008b). Considering a congested traffic condition, for instance, the maximum velocity of a car might be equal to the bicycle and walk modes. To address this issue, one solution is to extract more features from a GPS track (e.g., top five velocities rather than a single maximum velocity). However, if a lot of features are generated to cover more aspects of GPS trajectories, the challenge of applying an effective dimensionality reduction process needs to be met. Manual features are typically produced based on feature engineering, which is a concept upon biased engineering justification and commonsense knowledge of the real world for creating features and making patterns more visible for learning algorithms. Since the ultimate performance of machine learning algorithms is contingent on how much accurate the hand-crafted features are, the effectiveness proof of such features is not trivial. One way to address the above-mentioned issues is to exploit deep learning algorithms that are able to automatically and without any human interference extract multiple levels of data representations.

Indeed, the input layer (i.e., input features) in deep learning architectures is the raw object (e.g., images) rather than a set of hand-crafted features. The key role of deep learning techniques is to encode object's raw and low-level features (e.g., raw image pixels) to multiple levels of efficient and high-level features. This is the most salient attribute of deep learning algorithms that distinguishes them from classical machine learning algorithms. Thus, new representations of the raw data, which are more effective for the classification task, are generated by machines instead of humans. It should be noted that the learned features in the last layer play the same role as hand-crafted features, which are fed into activation functions (e.g., SVM or softmax) to compute class scores.

In this paper, we propose a Convolutional Neural Network (CNN) architecture to predict the transportation mode(s) used in an individual's trip from their raw GPS trajectories, in which modes are categorized into walk, bike, bus, driving, and train. The CNN is a type of deep learning techniques that has achieved great success in the fields of computer vision (Krizhevsky et al., 2012) and natural language processing (Kim, 2014). We envision to investigate the capability of CNNs in representation learning and transport mode classification of raw GPS data. Unlike other fields (e.g., image classification) that images are easily utilized as the input layer, the key challenge in this study is to structure raw GPS tracks into a format that is not only acceptable for CNN architectures but efficient enough to represent fundamental motion characteristics of a moving object. In our methodology, an instance comprises four channels of kinematic features including speed, acceleration, jerk, and bearing rate. Stacking these channels yields a standard arrangement for the CNN scheme that also describes people's motion characteristics. After pre-processing data and designing a suitable layout for each instance, we come up with an effective CNN architecture so as to attain state-of-the-art accuracy on the GPS dataset collected by the Microsoft GeoLife project (Zheng et al., 2008b).

The rest of this article is organized as follows. After reviewing related works in Section 2, we set out details of our framework in Section 3, including preparing the input layer, applying data pre-processing steps, explaining settings of CNN layers, and generating several CNN configurations for our application. In Section 4, we evaluate our proposed CNN architecture on the GeoLife trajectory dataset. Comparing our results with classical machine learning algorithms and previous research is also carried out in Section 4. Finally, we conclude the paper in Section 5.

2. Literature review

A large and growing body of literature has proposed numerous frameworks for inferring the commuters' transport mode based on various data sources including raw GPS trajectories (Zheng et al., 2008a; Endo et al., 2016; Xiao et al., 2017), mobile phone's accelerometers (Nick et al., 2010), and mobile phone's GSM data (Sohn et al., 2006). For performing the mode classification task, a wide range of traditional supervised mining algorithms have been applied, including rule-based methods, fuzzy logic, decision tree, Bayesian belief network, multilayer perceptron, and support vector machine (Wu et al., 2016). Furthermore, some of the mobility detection research has integrated multiple sources to ameliorate the classification quality. For example, Stenneth et al. (2011) exploited the GPS and GIS information for building their mode detection scheme while (Feng and Timmermans, 2013) combined GPS and accelerometer data to generate a model that outperforms GPS-only information. In addition to the GPS and accelerometer, the data from other mobile phone sensors such as gyroscope, rotation vector, and magnetometer have been deployed in distinguishing between different transportation modes (Jahangiri and Rakha, 2014; Eftekhari and Ghatee, 2016). Integrating such spatial and