Computers, Environment and Urban Systems 67 (2018) 124–131

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

# Physical activity classification in free-living conditions using smartphone accelerometer data and exploration of predicted results

Kangjae Lee[a,*], Mei-Po Kwan[b]

[a] Illinois Informatics Institute, University of Illinois at Urbana-Champaign, National Center for Supercomputing Applications, MC-257, 1205 W Clark St., Urbana, IL 61801, USA

[b] Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, 255 Computing Applications Building, MC-150, 605 E Springfield Ave., Champaign, IL 61820, USA

ARTICLE INFO

*Keywords:*
Physical activity classification
Machine learning
Accelerometer data
GPS trajectory
GIS

ABSTRACT

In recent decades, decreasing physical activity has emerged as one of the major issues affecting human health since people increasingly engaged in sedentary behavior in their homes and workplaces. In physical activity research, using GPS trajectories and advanced GIS methods has a potential for greatly enhancing our understanding of the association between objectively measured moderate and vigorous physical activity and physical and social environments. Relying only on objectively measured physical activity intensity, however, ignores the role of different places and types of physical activity on people's health outcomes. The aim of this study is to propose an approach to classifying physical activity in free-living conditions for physical activity research using published smartphone accelerometer data. Random forest and gradient boosting are used to predict jogging, walking, sitting, and standing. Generated training models based on the two classifiers are tested on accelerometer data collected from the smartphones of two subjects in free-living conditions. GPS trajectories with predicted physical activity labels are visually explored on a map to offer new insight on the assessment of the predicted results of daily activities and the identification of any difference in the results between random forest and gradient boosting. The findings of this study indicate that random forest and gradient boosting enable accurate physical activity classification in free-living conditions. GPS trajectories linked with predicted labels on a map assist the visual exploration of the erroneous prediction in daily activities including in-vehicle activities.

## 1. Introduction

In recent decades, decreasing physical activity (PA) has emerged as one of the major issues affecting human health since people increasingly engaged in sedentary behavior in their homes and workplaces. Moreover, obese people spend less time on moderate to vigorous physical activity (MVPA) when compared to non-obese people (Hagströmer, Troiano, Sjöström, & Berrigan, 2010). MVPA, such as brisk walking, bicycling, and jogging, contributes to reduced risk of physical and mental health problems (e.g., cardiovascular diseases, type II diabetes, obesity, depression, anxiety, and well-being) (Fox, 1999; Gordon-Larsen, Nelson, Page, & Popkin, 2006; Physical Activity Guidelines Advisory Committee, 2008; Wei, Gibbons, Kampert, Nichaman, & Blair, 2000). Metabolic syndrome is associated with a number of psychiatric disorders (Ho, Zhang, Mak, & Ho, 2014), and depression is a common comorbidity (Quek, Tam, Zhang, & Ho, 2017). Depression and obesity share a common pathological mechanism (Yang, De Xiang Liu, Pan, Ho, & Ho, 2016). PA and exercise lead to significant reduction of stress levels as compared to short-term pharmacological treatment (Lu et al., 2017). A number of scholars have attempted to identify factors in the physical and social environments that have significant positive or negative effects on people's PA. Many studies have examined the relationship between PA and specific characteristics of the built environment, including green spaces, based on neighborhood areas using geographic information systems (GIS) (Cohen et al., 2006; Coombes, Jones, & Hillsdon, 2010; McGinn, Evenson, Herring, Huston, & Rodriguez, 2007; Nagel, Carlson, Bosworth, & Michael, 2008; Saelens & Handy, 2008; Sallis et al., 2016).

In PA research, using GPS trajectories and advanced GIS methods has a potential for greatly enhancing our understanding of the association between objectively measured MVPA and physical and social environments (Browning & Lee, 2017). Through taking into account people's daily activities and travel, insights may be obtained to better inform policies or measures that seek to promote PA (Almanza, Jerrett, Dunton, Seto, & Pentz, 2012; Boruff, Nathan, & Nijënstein, 2012; Cooper et al., 2010; Helbich et al., 2016; Lachowycz, Jones, Page,

* Corresponding author.
*E-mail addresses:* klee171@illinois.edu (K. Lee), mpk654@gmail.com (M.-P. Kwan).

http://dx.doi.org/10.1016/j.compenvurbsys.2017.09.012
Received 15 April 2017; Received in revised form 25 September 2017; Accepted 30 September 2017
Available online 04 October 2017

Wheeler, & Cooper, 2012; Rodríguez et al., 2012; Troped, Wilson, Matthews, Cromley, & Melly, 2010). More important, recent studies using GPS trajectories confirm the importance of non-residential contexts (e.g., workplaces or locations for routine activities) in people's daily life as well as areas around their residential neighborhoods (Diez Roux & Mair, 2010; Inagami, Cohen, & Finch, 2007; Kwan, 2012a, 2012b; Perchoux, Chaix, Cummins, & Kestens, 2013). In most existing research, MVPA and sedentary behavior are determined using some thresholds based on the count per minute, an automatically calculated measure from commercial accelerometers (Berlin, Storti, & Brach, 2006; Frank, Schmid, Sallis, Chapman, & Saelens, 2005; Freedson, Melanson, & Sirard, 1998; Jones, Coombes, Griffin, & van Sluijs, 2009; Saelens, Sallis, Black, & Chen, 2003).

Relying only on the objectively measured intensity of PA, however, ignores the role of different places and types of PA on people's health outcomes. As a result, our knowledge about what types of PA were undertaken and what contextual characteristics are associated with healthy behaviors is limited when using intensity. In this context, Jankowska, Schipperijn, and Kerr (2015) highlighted the importance of understanding individual health behaviors over space and time related to PA and the need to go beyond using intensity. Thus, the specific types rather than the intensity of PA are critical to a better understanding of the association between PA and certain environmental influences, and accurate classification of PA types needs to be studied.

Many studies in the healthcare domain have been conducted to recognize different types of daily activities and PA using raw accelerometer data collected from subjects under controlled conditions (Anguita, Ghio, Oneto, Parra, & Reyes-Ortiz, 2012; Arif, Bilal, Kattan, & Ahamed, 2014; Kwapisz, Weiss, & Moore, 2011; Yin, Yang, & Pan, 2008; Zhang, McCullagh, Nugent, & Zheng, 2010; Zhu & Sheng, 2011). Mobile phone applications provide a low-cost technology, which allows clinicians to monitor PA of their patients without any technical knowledge (Zhang et al., 2014). The identification of different activity types enables the detection of abnormal behaviors when monitoring elderly people for their healthcare, the examination of the association between PA and its health effects on people, and the provision of feedback through mobile applications to encourage people to engage in PA. Further smartphone innovations are also helpful to care givers who are caring for individuals with dementia to improve not only their PA (Zhang et al., 2016) but also rehabilitation (Zhang, Yeo, & Ho, 2015). In the existing studies, machine learning techniques played an important role in building models based on a set of features derived from raw accelerometer data to predict various PA types. Machine learning is a branch of artificial intelligence and helps to predict outcomes after models/algorithms are trained using a large amount of input data. Features in machine learning are informative quantifiable attributes, derived from the input data, such as mean and standard deviation, used to determine different labels (or classes) within an acceptable range in models. The classification algorithms in existing studies, however, are mostly tested on accelerometer data collected under controlled situations. In other words, how daily activities and PA, which might include a variety of uncontrolled activities, can be represented by restricted types in the training accelerometer data needs to be understood. Therefore, a validation process needs to be performed to unveil how the accelerometer data collected in a laboratory setting provide convincing predicted results in various daily activities.

The aim of this study is to propose an approach to PA classification in free-living conditions for PA research using published smartphone accelerometer data. Free-living conditions refer to the natural everyday settings in people's daily lives in contrast to artificial laboratory conditions. Two supervised machine learning classifiers – random forest and gradient boosting – are used to generate training models based on publicly released accelerometer data for predicting different PA types and comparing their performance. The PA types identified by the proposed classification algorithm are jogging, walking, sedentary status, and standing. The performance of the generated predictive models is assessed in two different ways – 1) with an approach of cross-validation and 2) using test accelerometer data collected from a smartphone of one adult subject in free-living conditions. Because the published accelerometer data used to train models were collected under controlled conditions, the assessment of the learning models using data recorded in uncontrolled daily life is critical for this study. For more thorough examination of the models, a visual exploration of classified PA types over space and time is performed on a map using a set of GPS and accelerometer data collected from the smartphones of two subjects.

The process of building models developed in this study contributes to improving the performance of PA classification by highlighting practical strategies and considerations for collecting GPS and accelerometer data from human subjects. Further, it will be helpful to future studies that seek to advance the examination of the association between PA and environmental factors. The construction of learning models using publicly available accelerometer data with labels of PA types also will enable researchers to address the daunting challenge of requiring subjects in PA research to record the labels for every activity. Visualization of the classified types on GPS trajectories offers new insight into the assessment of the predicted results of daily activities and the identification of any difference in the results between the random forest and gradient boosting classifiers.

The sections in this paper are structured as follows: Section 2 summarizes existing studies on PA classification algorithms using accelerometer data. Section 3 describes the accelerometer data used in this study, preprocessing of the accelerometer data, and a classification algorithm using random forest and gradient classifiers taking into account extracted features from accelerometer data instances (samples). Section 4 demonstrates the performance of random forest and gradient boosting and the application of the developed classification algorithm to the accelerometer data collected from two subjects in free-living conditions. Lastly, discussion and conclusions of the findings in this study are presented in Section 5.

## 2. Related work

Due to the widespread use of smartphones, researchers are paying increasing attention to the sensing capabilities of smartphones. As one of the sensors in smartphones, tri-axial accelerometer records the accelerations of x, y, and z-axes, which may allow the recognition of different types of human activities. Many scholars in the public health and computer science domains have highlighted the potential of the built-in accelerometer sensor in smartphones to recognize different types of PA using machine learning techniques (Anguita et al., 2012; Arif et al., 2014; Kwapisz et al., 2011; Zhang et al., 2010). For instance, Zhang et al. (2010) used support vector machine (SVM) to recognize six PA types including walking, standing, and sitting. Three-axis accelerometer data were recorded by 10 participants using smartphones worn on their waists on the left side. Kwapisz et al. (2011) developed a system to recognize six PA types by analyzing collected accelerometer data through smartphones of participants. The recognition process was performed using 43 generated features extracted from raw accelerometer data and tested using three different machine learning models. Among the three machine learning classifiers, the multilayer perceptron showed 91.7% predictive accuracy. Arif et al. (2014) used the same accelerometer dataset of Kwapisz et al. (2011) to classify the same six PA types. Optimal features derived from raw accelerometer taking into account correlation coefficients helped to increase the accuracy of prediction to 97% using K nearest neighborhood. To achieve a higher level of predictive accuracy, a combination of the accelerometer and other built-in sensors in smartphones (e.g., gyroscope, magnetometer) was also explored (Anguita et al., 2012; Shoaib, Scholten, & Havinga, 2013). With the recognition technique, PA monitoring systems were developed to detect abnormal activities of elderly people or people in need of assistance for healthcare purposes (Anguita et al., 2012; Ketabdar & Lyra, 2010). Further, the monitoring systems could encourage people to engage in regular PA to achieve recommended levels (Zhang et al., 2010). Apart from people's healthcare, abnormal activity detection based on activity recognition algorithms was also exploited to identify terrorist activities for security reasons (Yin et al., 2008).

In this study, a PA classification algorithm is developed to identify jogging, walking, sedentary status, and standing using publicly released wireless sensor data mining (WISDM)'s accelerometer data version 1.1 (Kwapisz et al., 2011), which is described in Section 3.1. Compared to previous studies, this study seeks to advance PA research through not only developing a classification algorithm but also assessing its application to accelerometer data collected from people in free-living conditions (using raw accelerometer and GPS data collected from two subjects). Only a few studies (e.g., Nguyen, Moore, & McCowan, 2007) have attempted to apply the capabilities of machine learning to recognize free-living daily activities. However, the methods used in existing research are limited to using unsupervised machine learning methods, like clustering, which only differentiate between two or more groups of data objects (e.g., free-living daily activities vs. abnormal activities) without exact labels as results regarding specific PA types. Hence, in this study, the role of the supervised machine learning models – random forest and gradient boosting – is vital to predicting specific PA types and to identify any difference in daily activities under free-living conditions represented as predicted results from the classifiers.

## 3. A physical activity classification algorithm

Because the WISDM data used in this study are labeled with six PA types, some types (e.g., going upstairs) are merged together with similar PA types (e.g., walking) (Section 3.1). The raw form of accelerometer data needs to be grouped into an analytic unit, called example, for feature extraction (Section 3.2). In this study, a set of 200 consecutive raw instances (acceleration records during 10 s) of the identical PA-type label forms one example (subset of accelerometer data), which showed the best predictive accuracy in the research by Kwapisz et al. (2011). Next, to predict different PA types using machine learning classifiers, various informative features are derived from the grouped unit (Section 3.3). The generated features are then, used to classify four PA types using tree-based multiclass classifiers (Section 3.4).

### 3.1. Description of the training accelerometer data

The WISDM accelerometer data were collected from 36 persons using Android smartphones under laboratory conditions with controlled settings (Table 1). Each of the subjects was asked to carry a smartphone in their pant pocket, which recorded accelerations at 20 hertz (Hz) (20 samples/s) with a timestamp. The WISDM data have more than one million of instances, and each instance includes not only x, y, and z-acceleration values but also a label of PA type. Among a total of 6 different types, walking and jogging account for most of the labels in the dataset – 39% and 31% of the number of instances, respectively – whereas standing only accounts for 4%. In this study, since going upstairs and downstairs account for a small portion of the individuals' daily PA and resemble walking in terms of body motion, these activity types are merged into walking.

**Table 1**
Description of wireless sensor data mining (WISDM)'s accelerometer data.

| Data | Number of instances | Attributes | Physical activity type | Sampling rate |
|---|---|---|---|---|
| Wireless sensor data mining (WISDM)'s accelerometer dataset v1.1 | 1,098,207 | Person ID, physical activity mode, timestamp, and acceleration (x, y, z) | Walking (39%[a]), jogging (31%), upstairs (11%), downstairs (9%), sitting (6%), and standing (4%) | 20 Hz (20 samples/s) |

[a] Percentages of each PA label are calculated based on the total number of instances.

### 3.2. Preprocessing of accelerometer data

Preprocessing of the WISDM accelerometer data is performed to select valid accelerometer instances and to define and generate an analytic unit for feature extraction in the next stage. Firstly, x, y and z-acceleration values are smoothed using a low pass filter. The low pass filter helps to clean up the acceleration signals and minimize the effect of noises. Secondly, all the raw instances (samples) of the accelerometer data are grouped into examples to compute many features (e.g., average of x acceleration, a standard deviation of y acceleration). Lastly, a set of features for each example are calculated by using a moving window. Particularly, a 50% overlap moving window is applied to detect more meaningful patterns and improve predictive accuracy taking into account some instances in the windows in the previous and next turns (Bao & Intille, 2004; Ravi, Dandekar, Mysore, & Littman, 2005). An illustration of the moving windows on 1000 points of the acceleration in the x-direction is presented in Fig. 1. One moving window covers one example (200 samples) and shifts forward (to the right in Fig. 1) making a 50% overlap with the previous window. That means, 100 samples in the previous turn are taken into account to calculate a set of features in the current window. Thus, 9 examples with a set of calculated features are generated in the illustration below.

### 3.3. Feature extraction

A set of 59 features are extracted from each example of 200 instances by mathematical and statistical calculations in this study, and the details of all features are described in Table 2. Most features are the same and expanded features as those identified by Kwapisz et al. (2011) and Bao and Intille (2004), which are already shown to be important for the classification of PA types using accelerometer data. Among the features, the mean dominant frequency and mean energy of frequency are calculated by using Fast Fourier Transform (FFT) (Welch, 1967). This study especially adds two more kinds of features – 1) standard deviation of the total acceleration and 2) min-max mean value of the total acceleration – besides the features used in the two previous studies. The magnitude of the total acceleration is calculated by the square root of the sum of squared acceleration of three axes as in Eq. (1):

$$\text{Total acceleration} = \sqrt{x^2 + y^2 + z^2}, \tag{1}$$

where x, y, and z stand for raw acceleration of each x, y and z-axis. In addition to using the time interval between local peaks, counts of local peaks are also calculated for each example to consider the different frequencies of peak occurrence that may vary between different types of PA.

### 3.4. Classification

Multiclass classification models are capable of identifying more than two different classes of discrete attributes, groups, characteristics, or objects that researchers want to predict based on extracted features from a given dataset. Among multiclass classification machine learning models, random forest (Breiman, 2001) and gradient boosting (Friedman, 2001) are used to identify four PA types based on the 59 features extracted from the published accelerometer data. Random forest grows multiple trees and classifies labels based on the votes of all the trees. Among all features, a subset of features are randomly selected

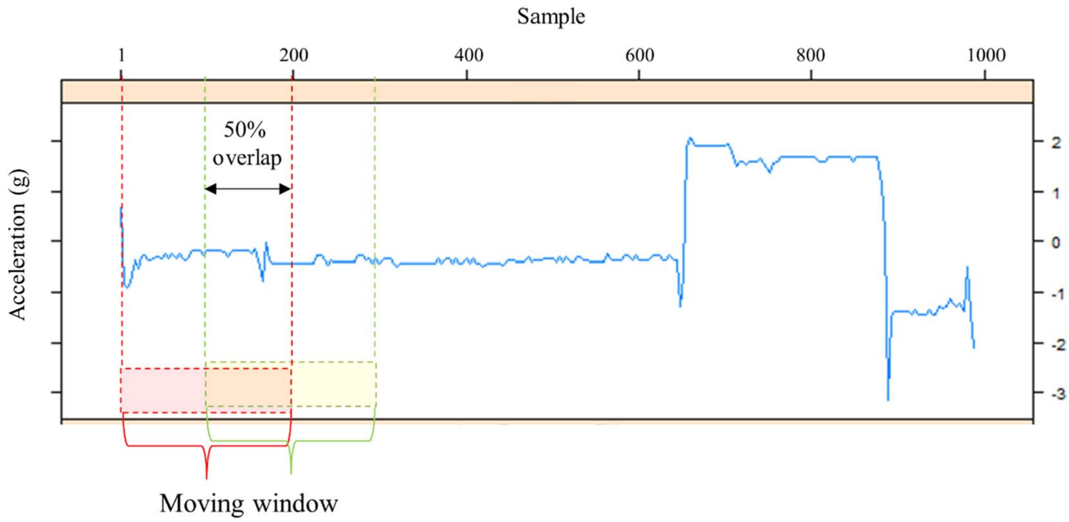


**Fig. 1.** 50% overlap moving window on 1000 points of the acceleration in the x direction.

**Table 2**
All 59 features extracted from each accelerometer example.

| Feature | Description |
|---|---|
| Mean | Mean of three axes and total acceleration (4 features) (Kwapisz et al., 2011) |
| Standard deviation | Standard deviation of three axes and total acceleration (4 features) (Kwapisz et al., 2011) |
| Binned range | Number of examples falling within each 10 bin of three axes (30 features) (Kwapisz et al., 2011) |
| Min-max mean | Mean of the absolute different between minimum and maximum values of three axes and total acceleration (4 features) (Kwapisz et al., 2011) |
| Time interval between local peaks | Mean time interval between two consecutive local peaks (3 features) (Kwapisz et al., 2011) |
| Counts of local peaks | Number of local peaks (3 features) |
| Correlation | Correlation among three axes (3 features) (Bao & Intille, 2004) |
| Mean dominant frequency | Mean of the first 3 dominant frequencies of three axes and total acceleration using Fast Fourier Transform (frequency with highest amplitude) (4 features) (Bao & Intille, 2004) |
| Mean energy of frequency | Mean value of energy of the frequencies for three axes and total acceleration (4 features) (Bao & Intille, 2004) |

for each node, and the best split on the subset of features is chosen to split node. On the other hand, gradient boosting as a forward stage-wise optimization algorithm uses votes of each weak classifier, which is learned at every iteration, to generate a strong classifier. Gradient boosting uses regression tree models as weak classifiers and generates a strong model based on the notion of gradients in a way that a loss function is minimized. These two powerful machine learning classifiers based on tree models are expected to achieve high predictive accuracy, and the results obtained using them are compared and evaluated. The comparison and evaluation of the results are described in Section 4.

R statistical computing software provides some packages based on the random forest and gradient boosting framework (R Development Core Team, 2008). The packages used in this study are 'randomForest' and 'xgboost' for the random forest and gradient boosting classifiers (Liaw & Wiener, 2002; Chen, He, & Benesty, 2016). The 'xgboost' particularly supports parallel computation and shows outstanding performance in terms of computation time when compared to other gradient boosting packages in R.

## 4. Classification results

### 4.1. Classification results using WISDM accelerometer data

10-fold cross-validation evaluated the performance of the learning models generated using random forest and gradient boosting. The number of the total examples is 10,243, and among them, 10% is randomly sampled to make test data in each iteration for 10-fold cross-validation, taking into account the original proportion of each PA type. Random forest and gradient boosting classifiers are run and predictive accuracy is 99.03% and 99.22 respectively. Table 3 shows the predictive accuracy when random forest is used for the classification, where all types achieve very high accuracy of over 98%. High predictive accuracy is also achieved when gradient boosting is used (Table 4). In the confusion matrix of the random forest classifier, walking and jogging have the lowest classification accuracy. 46 walking examples and 48 jogging examples are incorrectly predicted as jogging and walking respectively. The use of a 50% overlap window helps to increase the prediction accuracy of both random forest and gradient boosting by 0.54% and 0.44% respectively, as shown in Table 5.

### 4.2. Classification results using accelerometer data collected in free-living conditions

The two learning models generated are tested using accelerometer data collected from one adult subject in free-living conditions to validate the applicability of the generated learning models (Table 6). The accelerometer data have 49,016 instances recorded at 20 Hz using an LG G3 smartphone and have six PA types same as those of the WISDM data, with sufficient recording times. Walking up and down stairs are merged into walking to make the labels the same as the labels used in the learning models. Because both random forest and gradient boosting generally provide different predictions for every run with the same

**Table 3**
Confusion matrix of physical activity classification using random forest.

| | | Actual class | | | |
|---|---|---|---|---|---|
| | | Jogging | Walking | Sitting | Standing |
| Predicted class | Jogging | 3062 | 48 | 0 | 1 |
| | Walking | 46 | 6108 | 1 | 1 |
| | Sitting | 0 | 0 | 536 | 0 |
| | Standing | 0 | 0 | 2 | 438 |
| Accuracy (%) | | 98.52 | 99.22 | 99.44 | 99.55 |

**Table 4**
Confusion matrix of physical activity classification using gradient boosting.

| | | Actual class | | | |
|---|---|---|---|---|---|
| | | Jogging | Walking | Sitting | Standing |
| Predicted class | Jogging | 3070 | 34 | 0 | 1 |
| | Walking | 37 | 6122 | 2 | 1 |
| | Sitting | 1 | 0 | 535 | 2 |
| | Standing | 0 | 0 | 2 | 436 |
| Accuracy (%) | | 98.78 | 99.45 | 99.26 | 99.10 |

**Table 5**
Comparison of performance between random forest and gradient boosting with no overlap or a 50% overlap window.

| No overlap window | | 0.5% overlap window | |
|---|---|---|---|
| Random forest | Gradient boosting | Random forest | Gradient boosting |
| 98.49% | 98.78% | 99.03% | 99.22% |

**Table 6**
Description of accelerometer data instances collected under free-living conditions to test the learning models.

| Number of samples | Physical activity type | Sampling rate |
|---|---|---|
| 49,016 | Walking (upstairs and downstairs: 12%, walking: 29%), jogging (17%), sitting (15%), and standing (27%) | 20 Hz (20 samples / s) |

data, averaged predictive accuracy is calculated for all 200 iterations of the two classifiers.

Confusion matrices of one predicted result obtained with the random forest and gradient boosting classifiers using the accelerometer data are presented in Tables 7 and 8. The averaged predictive accuracy of the random forest classifier in the 200 iterations is 95.10%, which is almost 4% lower than the accuracy of its training model, whereas gradient boosting has a 99.10% predictive accuracy. The confusion matrix of random forest (Table 7) indicates that almost 9% of the walking examples are wrongly classified as jogging. It is caused by some instances recorded during brisk walking at a fast pace with the walking label, which is confirmed in Section 4.2 based on the exploration of GPS trajectories coupled with classified types. On the other hand, in the classification using gradient boosting, walking examples including walking up and down stairs are classified with high accuracy (97.80%). For sitting and standing, both random forest and gradient boosting achieve perfect classification accuracy.

### 4.3. Exploration of classified PA types with GPS tracks in free-living conditions

Generated learning models using the two classifiers are evaluated using accelerometer and GPS data collected in free-living conditions. The evaluation seeks to show how classification results from the classification algorithm using random forest or gradient boosting are represented in various daily activities in free-living conditions by visualizing GPS trajectories with predicted PA labels. Adult subject 1 and 2 were given LG G3 and/or Samsung Galaxy Alpha to record both GPS tracks and accelerometer data during 28 and 8 days, respectively (Table 9). Subject 1 used two different devices to identify whether there is any difference in predicted types using accelerometer data collected from the two smartphones. The GPS tracks were recorded at a 1-second interval together with a measure of positional accuracy called horizontal dilution of precision (HDOP), which represents the quality of a horizontal position of a GPS point. In this study, HDOP $< 5$ is used as a criterion for removing GPS points with low accuracy, and removed GPS points are replaced with re-estimated points using a linear interpolation method based on the coordinates of the GPS points with high location accuracy. Then, predicted types are combined with the GPS points based on the timestamp that both accelerometer data and GPS tracks have.

Fig. 2 demonstrates how accurately free-living activities are predicted by random forest and gradient boosting classifiers for the four PA types. GPS trajectories coupled with predicted types are visualized on OpenStreetMap to assess the prediction results with movement patterns on GPS tracks and geographic contexts. Most jogging and walking behaviors are correctly classified by random forest and gradient boosting. Characteristics of some of the walking examples that are wrongly classified as jogging are visually examined. Sitting and standing are also accurately predicted, and they occur in buildings, especially at work places and homes and near bus stops (not shown on the map) respectively. The series of predicted sitting postures indoors is represented as a contrived straight line because indoor GPS trajectories have low accuracy due to signal lose caused by barriers and thus, only a few highly accurate GPS points are used to interpolate GPS points. In other words, most indoor GPS points are likely to have low positional accuracy and are in turn replaced by estimated GPS points using a linear interpolation method based on a few highly accurate GPS points. Hence, the recorded indoor GPS tracks, which represent sitting in Fig. 2, do not mean that the subject was moving in sitting postures; so does the indoor shopping activity shown in Fig. 2. Even though it looks visually erroneous, the concentration of points on buildings is really helpful for the accurate estimation of the effect of environments using buffer analysis or point density-based analysis considering space and time dimensions.

**Table 7**
Confusion matrix in one of 200 iterations in physical activity classification using random forest.

| | | Actual class | | | |
|---|---|---|---|---|---|
| | | Jogging | Walking | Sitting | Standing |
| Predicted class | Jogging | 38 | 8 | 0 | 0 |
| | Walking | 1 | 83 | 0 | 0 |
| | Sitting | 0 | 0 | 33 | 0 |
| | Standing | 0 | 0 | 0 | 60 |
| Accuracy (%) | | 97.44 | 91.21 | 100 | 100 |

**Table 8**
Confusion matrix of one of 200 iterations in physical activity classification using gradient boosting.

| | | Actual class | | | |
|---|---|---|---|---|---|
| | | Jogging | Walking | Sitting | Standing |
| Predicted class | Jogging | 39 | 2 | 0 | 0 |
| | Walking | 0 | 89 | 0 | 0 |
| | Sitting | 0 | 0 | 33 | 0 |
| | Standing | 0 | 0 | 0 | 60 |
| Accuracy (%) | | 100 | 97.80 | 100 | 100 |

**Table 9**
Description of GPS and accelerometer data collected under free-living conditions.

| Subject ID | Device | Sampling rate | | Recording time |
|---|---|---|---|---|
| | | GPS | Accelerometer | |
| 1 | LG G3, Samsung Galaxy Alpha | 1 s | 20 Hz (20 instances/s) | 28 days |
| 2 | Samsung Galaxy Alpha | | | 8 days |

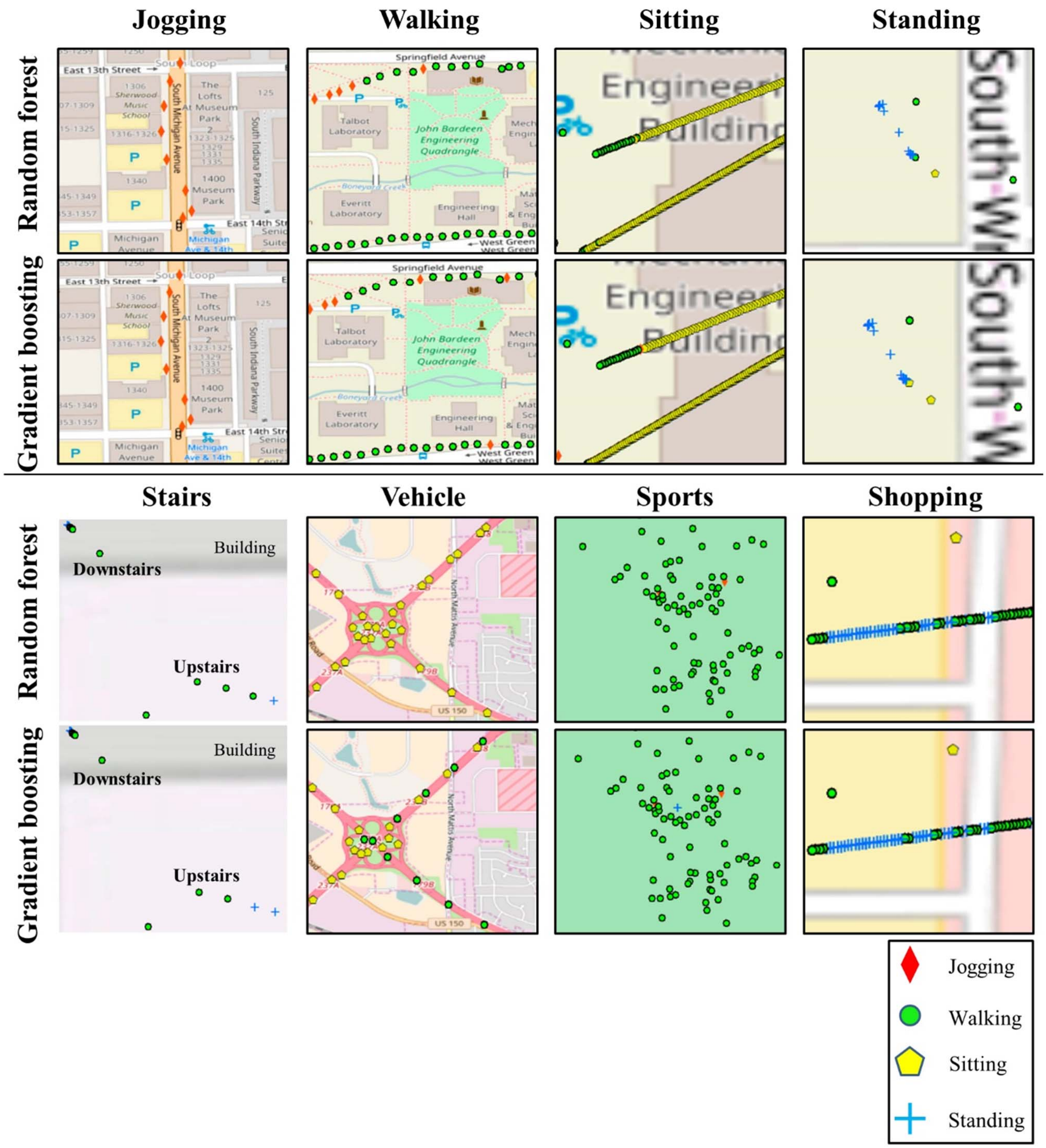


**Fig. 2.** Comparison of classified physical activity types using gradient boosting and random forest along with GPS tracks.

Upstairs and downstairs are successfully predicted as walking, and only one GPS point has a differently classified label in upstairs between random forest and gradient boosting. GPS points for upstairs are located outside the building because of the erroneous GPS signal indoors.

In addition to the four basic PA types, other kinds of movement behaviors are explored to assess the predictive accuracy of the two classifiers for different daily activities. Both random forest and gradient boosting yield similar and plausible results for sports and shopping activities. Many points with a walking status are incorporated as parts of the sports activities classified in the process whereas standing is the most commonly predicted type among the shopping activities. Note that gradient boosting predictes slightly more standing examples in the sports and shopping activities when compared to those predicted by random forest. On the other hand, only random forest shows consistent results in the prediction of the subjects' postures when they were in a vehicle (e.g., a car or a bus). Because the subjects were sitting in a car or a bus, most of the activities that occurred inside a vehicle are sitting, and random forest tends to correctly classify the subjects' status in vehicles as sitting. Gradient boosting, however, shows inconsistent results for in-vehicle activities as shown in Fig. 2. Some of the GPS points on highways are incorrectly classified as walking by gradient boosting.

## 5. Discussion and conclusions

In this study, an approach to classifying PA in free-living conditions was proposed using random forest and gradient boosting, and the predictive accuracy of the two classifiers was assessed. These two classifiers achieved highly accurate classification with respect to jogging, walking, sitting, and standing activities in both controlled and free-living conditions. Particularly, the high accuracy obtained through the algorithm developed using accelerometer data collected under laboratory conditions was found to be valid in the classification of the accelerometer data collected in free-living conditions without segmenting accelerometer records based on the same type. The set of 59 features extracted from instances of the WISDM accelerometer data contributed to the high prediction accuracy of over 99% on the 10-fold cross validation and over 95% on the accelerometer data collected under free-living conditions with random forest and gradient boosting. In addition, the use of a 50% overlap window contributed to the slight increase in predictive accuracy.

In the case of free-living conditions, random forest and gradient boosting classified all activities with a series of plausible PA types in visualization even though the accelerometer data collected from the two subjects in free-living conditions include instances of uncontrolled movement in different surroundings and different kinds of activities. For example, sitting in a vehicle and sitting and standing in a building were classified correctly based on exploratory analysis, and the uncontrolled movement during sports was interpreted as a combination of postures that were characteristic of the sports. Particularly, random forest showed more consistent results for in-vehicle activities whereas gradient boosting was sensitive to the subjects' postures in moving vehicles. Further, GPS trajectories linked with the predicted labels on a map assisted the exploration of the ambiguity between fast walking and jogging. The ambiguity between brisk walking and walking at a normal pace was one crucial factor that significantly affected the performance on the accelerometer data. However, as long as brisk walking is classified as jogging not as standing or sitting, researchers in the PA domain may be able to reclassify it as walking based on its irregular presence in the middle of a walking trip or the speed derived from GPS tracks linked to predicted types if needed.

Regarding smartphone data collection, no matter which side of the smartphone faced when they were in the subjects' pant pocket, both random forest and gradient boosting showed consistent results overall. Nevertheless, if the smartphone was carried in other pockets, like a jacket pocket, the collected accelerometer data from the smartphones were not be able to predict types correctly. In fact, we obtained ludicrous results in classifying them when analyzing the accelerometer data collected from a third subject who carried his/her smartphone in a jacket pocket inadvertently. This indicates that when subjects' smartphones are out of their pant pockets for phone calls or playing games, PA types are likely to be incorrectly classified, and thus such use of smartphones needs to be restricted during the data collection process. In addition, the use of a low-priced smartphone hampered the accurate accelerometer data collection. The subject 2 was given a low-cost smartphone, LG Realm, at first. However, accelerometer data collected from the smartphone during one week showed erratic acceleration. We concluded that the low resolution of a built-in accelerometer sensor in a cheap smartphone is likely to collect low quality of accelerometer data, which will cause incorrect classification and thus, the selection of smartphones for accelerometer data collection is critical.

The proposed approach, however, has some limitations that need to be addressed in future research. First, post-processing is needed to handle fast walking trips in the middle of walking trips. In terms of the intensity of PA, fast walking is a kind of moderate exercise, and people consume more calories when they engage in jogging than walking. One option for post-processing is to reclassify brisk walking trips identified as jogging into walking based on a commonly used cut-off point on the speed of corresponding GPS points (e.g., 5 mph) and duration of the trips. Second, collecting supplementary data using daily activity diaries should be considered in future research. Because the developed classification algorithm cannot classify all different kinds of sports, especially with extreme upper body movement (Berlin et al., 2006), an effective means to collect what kinds of exercises subjects have performed needs to be considered. Third, transportation modes including biking and subways need to be classified in addition to the four types. Detection of transportation modes is really important in PA research because there is a correlation between high utilization of motorized transportation modes for commuting and the level of PA of young adults (Gordon-Larsen, Nelson, & Beam, 2005). Last, environments may have less impact on people's PA when they are in a bus or car, when compared to walking or running. Many studies in the transportation domain have already been conducted out to automatically recognize different transportation modes using GPS trajectories with/without accelerometer data (Ellis et al., 2014; Feng & Timmermans, 2013; Gong, Chen, Bialostozky, & Lawson, 2012; Moiseeva, Jessurun, & Timmermans, 2010; Zheng, Chen, Li, Xie, & Ma, 2010; Zhou, 2014). Thus, the methods proposed in these previous studies will be helpful for identifying transportation modes in PA research.

## References


Almanza, E., Jerrett, M., Dunton, G., Seto, E., & Pentz, M. A. (2012). A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data. *Health & Place, 18*(1), 46–54.

Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2012). Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. *International Workshop on Ambient Assisted Living* (pp. 216–223). Springer Berlin Heidelberg December.

Arif, M., Bilal, M., Kattan, A., & Ahamed, S. I. (2014). Better physical activity classification using smartphone acceleration sensor. *Journal of Medical Systems, 38*(9), 1–10.

Bao, L., & Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. *International conference on pervasive computing* (pp. 1–17). Springer Berlin Heidelberg April.

Berlin, J. E., Storti, K. L., & Brach, J. S. (2006). Using activity monitors to measure physical activity in free-living conditions. *Physical Therapy, 86*(8), 1137.

Boruff, B. J., Nathan, A., & Nijënstein, S. (2012). Using GPS technology to (re)-examine operational definitions of 'neighbourhood'in place-based health research. *International Journal of Health Geographics, 11*(1), 1.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Browning, M., & Lee, K. (2017). Within what distance does "Greenness" best predict physical health? A systematic review of articles with GIS buffer analyses across the lifespan. *International Journal of Environmental Research and Public Health, 14*(7), 675.

Chen, T., He, T., & Benesty, M. (2016). xgboost: Extreme gradient boosting. R package version 0.4–4. https://CRAN.R-project.org/package=xgboost.

Cohen, D. A., Ashwood, J. S., Scott, M. M., Overton, A., Evenson, K. R., Staten, L. K., ... Catellier, D. (2006). Public parks and physical activity among adolescent girls. *Pediatrics, 118*(5), 1381–1389.

Coombes, E., Jones, A. P., & Hillsdon, M. (2010). The relationship of physical activity and overweight to objectively measured green space accessibility and use. *Social Science & Medicine, 70*(6), 816–822.

Cooper, A. R., Page, A. S., Wheeler, B. W., Griew, P., Davis, L., Hillsdon, M., & Jago, R. (2010). Mapping the walk to school using accelerometry combined with a global positioning system. *American Journal of Preventive Medicine, 38*(2), 178–183.

Diez Roux, A. V., & Mair, C. (2010). Neighborhoods and health. *Annals of the New York Academy of Sciences, 1186*(1), 125–145.

Ellis, K., Godbole, S., Marshall, S., Lanckriet, G., Staudenmayer, J., & Kerr, J. (2014). Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms. *Public Health, 2*, 39–46.

Feng, T., & Timmermans, H. J. (2013). Transportation mode recognition using GPS and accelerometer data. *Transportation Research Part C: Emerging Technologies, 37*, 118–130.

Fox, K. R. (1999). The influence of physical activity on mental well-being. *Public Health Nutrition, 2*(3a), 411–418.

Frank, L. D., Schmid, T. L., Sallis, J. F., Chapman, J., & Saelens, B. E. (2005). Linking objectively measured physical activity with objectively measured urban form: Findings from SMARTRAQ. *American Journal of Preventive Medicine, 28*(2), 117–125.

Freedson, P. S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise, 30*(5), 777–781.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics,* 1189–1232.

Gong, H., Chen, C., Bialostozky, E., & Lawson, C. T. (2012). A GPS/GIS method for travel mode detection in New York City. *Computers, Environment and Urban Systems, 36*(2), 131–139.

Gordon-Larsen, P., Nelson, M. C., & Beam, K. (2005). Associations among active transportation, physical activity, and weight status in young adults. *Obesity Research, 13*(5), 868–875.

Gordon-Larsen, P., Nelson, M. C., Page, P., & Popkin, B. M. (2006). Inequality in the built

environment underlies key health disparities in physical activity and obesity. *Pediatrics, 117*(2), 417–424.
Hagströmer, M., Troiano, R. P., Sjöström, M., & Berrigan, D. (2010). Levels and patterns of objectively assessed physical activity—A comparison between Sweden and the United States. *American Journal of Epidemiology, 171*(10), 1055–1064.
Helbich, M., van Emmichoven, M. J. Z., Dijst, M. J., Kwan, M. P., Pierik, F. H., & de Vries, S. I. (2016). Natural and built environmental exposures on children's active school travel: A Dutch global positioning system-based cross-sectional study. *Health & Place, 39*, 101–109.
Ho, C. S., Zhang, M. W., Mak, A., & Ho, R. C. (2014). Metabolic syndrome in psychiatry: Advances in understanding and management. *Advances in Psychiatric Treatment, 20*(2), 101–112.
Inagami, S., Cohen, D. A., & Finch, B. K. (2007). Non-residential neighborhood exposures suppress neighborhood effects on self-rated health. *Social Science & Medicine, 65*(8), 1779–1791.
Jankowska, M. M., Schipperijn, J., & Kerr, J. (2015). A framework for using GPS data in physical activity and sedentary behavior studies. *Exercise and Sport Sciences Reviews, 43*(1), 48.
Jones, A. P., Coombes, E. G., Griffin, S. J., & van Sluijs, E. M. (2009). Environmental supportiveness for physical activity in English schoolchildren: A study using Global Positioning Systems. *International Journal of Behavioral Nutrition and Physical Activity, 6*(1), 1.
Ketabdar, H., & Lyra, M. (2010). System and methodology for using mobile phones in live remote monitoring of physical activities. *IEEE International Symposium on Technology and Society* (pp. 350–356). June.
Kwan, M. P. (2012a). The uncertain geographic context problem. *Annals of the Association of American Geographers, 102*(5), 958–968.
Kwan, M.-P. (2012b). How GIS can help address the uncertain geographic context problem in social science research. *Annals of GIS, 18*(4), 245–255.
Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter, 12*(2), 74–82.
Lachowycz, K., Jones, A. P., Page, A. S., Wheeler, B. W., & Cooper, A. R. (2012). What can global positioning systems tell us about the contribution of different types of urban greenspace to children's physical activity? *Health & Place, 18*(3), 586–594.
Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News, 2*(3), 18–22.
Lu, C. Y., Liu, D. X., Jiang, H., Pan, F., Ho, C. S., & Ho, R. (2017). Effects of traumatic stress induced in the juvenile period on the expression of gamma-aminobutyric acid receptor type A subunits in adult rat brain. *Neural Plasticity,* 2017.
McGinn, A. P., Evenson, K. R., Herring, A. H., Huston, S. L., & Rodriguez, D. A. (2007). Exploring associations between physical activity and perceived and objective measures of the built environment. *Journal of Urban Health, 84*(2), 162–184.
Moiseeva, A., Jessurun, J., & Timmermans, H. (2010). Semiautomatic imputation of activity travel diaries: Use of global positioning system traces, prompted recall, and context-sensitive learning algorithms. *Transportation Research Record: Journal of the Transportation Research Board, 2183*, 60–68.
Nagel, C. L., Carlson, N. E., Bosworth, M., & Michael, Y. L. (2008). The relation between neighborhood built environment and walking activity among older adults. *American Journal of Epidemiology, 168*(4), 461–468.
Nguyen, A., Moore, D., & McCowan, I. (2007). Unsupervised clustering of free-living human activities using ambulatory accelerometry. *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4895–4898). August.
Perchoux, C., Chaix, B., Cummins, S., & Kestens, Y. (2013). Conceptualization and measurement of environmental exposure in epidemiology: Accounting for activity space related to daily mobility. *Health & Place, 21*, 86–93.
Physical Activities Guidelines Advisory Committee (2008). *Physical activity guidelines advisory committee report.* Washington (DC): US Department of Health and Human Services.
Quek, Y. H., Tam, W. W., Zhang, M. W., & Ho, R. (2017). Exploring the association between childhood and adolescent obesity and depression: A meta-analysis. *Obesity Reviews, 18*(7), 742–754.
R Development Core Team (2008). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0 http://www.R-project.org.
Ravi, N., Dandekar, N., Mysore, P., & Littman, M. L. (2005). Activity recognition from accelerometer data. *Association for the advancement of artificial intelligence* (pp. 1541–1546). (5).
Rodríguez, D. A., Cho, G. H., Evenson, K. R., Conway, T. L., Cohen, D., Ghosh-Dastidar, B., & Lytle, L. A. (2012). Out and about: association of the built environment with physical activity behaviors of adolescent females. *Health & Place, 18*(1), 55–62.
Saelens, B. E., & Handy, S. L. (2008). Built environment correlates of walking: A review. *Medicine and Science in Sports and Exercise, 40*(7 Suppl), S550.
Saelens, B. E., Sallis, J. F., Black, J. B., & Chen, D. (2003). Neighborhood-based differences in physical activity: An environment scale evaluation. *American Journal of Public Health, 93*(9), 1552–1558.
Sallis, J. F., Cerin, E., Conway, T. L., Adams, M. A., Frank, L. D., Pratt, M., Salvo, D., Schipperijn, J., Smith, G., Cain, K. L., et al. (2016). Physical activity in relation to urban environments in 14 cities worldwide: A cross-sectional study. *The Lancet, 387*(10034), 2207–2217.
Shoaib, M., Scholten, H., & Havinga, P. J. (2013). Towards physical activity recognition using smartphone sensors. *Ubiquitous Intelligence and Computing, 2013 IEEE 10th International Conference on and 10th International Conference on Autonomic and Trusted Computing (UIC/ATC)* (pp. 80–87). December.
Troped, P. J., Wilson, J. S., Matthews, C. E., Cromley, E. K., & Melly, S. J. (2010). The built environment and location-based physical activity. *American Journal of Preventive Medicine, 38*(4), 429–438.
Wei, M., Gibbons, L. W., Kampert, J. B., Nichaman, M. Z., & Blair, S. N. (2000). Low cardiorespiratory fitness and physical inactivity as predictors of mortality in men with type 2 diabetes. *Annals of Internal Medicine, 132*(8), 605–611.
Welch, P. D. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics, 15*(2), 70–73.
Yang, J. L., De Xiang Liu, H. J., Pan, F., Ho, C. S., & Ho, R. C. (2016). The effects of high-fat-diet combined with chronic unpredictable mild stress on depression-like behavior and leptin/leprb in male rats. *Scientific Reports, 6*.
Yin, J., Yang, Q., & Pan, J. J. (2008). Sensor-based abnormal human-activity detection. *IEEE Transactions on Knowledge and Data Engineering, 20*(8), 1082–1090.
Zhang, M., Cheow, E., Ho, C. S., Ng, B. Y., Ho, R., & Cheok, C. C. S. (2014). Application of low-cost methodologies for mobile phone app development. *JMIR mHealth and uHealth, 2*(4).
Zhang, M. W., Chan, S., Wynne, O., Jeong, S., Hunter, S., Wilson, A., & Ho, R. (2016). Conceptualization of an evidence-based smartphone innovation for caregivers and persons living with dementia. *Technology and Health Care, 24*(5), 769–773.
Zhang, M. W., Yeo, L. L., & Ho, R. C. (2015). Harnessing smartphone technologies for stroke care, rehabilitation and beyond. *BMJ Innovations, 1*(4), 145–150.
Zhang, S., McCullagh, P., Nugent, C., & Zheng, H. (2010). Activity monitoring using a smart phone's accelerometer with hierarchical classification. *Intelligent Environments (IE), 2010 Sixth International Conference on IEEE,* 158–163.
Zheng, Y., Chen, Y., Li, Q., Xie, X., & Ma, W. Y. (2010). Understanding transportation modes based on GPS data for web applications. *ACM Transactions on the Web (TWEB), 4*(1), 1.
Zhou, X. (2014). Investigating the association between the built environment and active travel of young adults using location based technology (Doctoral dissertation). Retrieved from IDEALS.
Zhu, C., & Sheng, W. (2011). Motion-and location-based online human daily activity recognition. *Pervasive and Mobile Computing, 7*(2), 256–269.