2020

Smartphone Sensor-based Pedestrian Activity Recognition for P2V Communication and Warning System

Dhrubo Hasan Chowdhury
University of Central Florida

SMARTPHONE SENSOR-BASED PEDESTRIAN ACTIVITY RECOGNITION FOR P2V COMMUNICATION AND WARNING SYSTEM

by

DHRUBO HASAN CHOWDHURY
B.Sc. Bangladesh University of Engineering and Technology, 2017

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Civil, Environmental and Construction Engineering
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Summer Term
2020

Major Professor: M. Abdel-Aty

`

`

ABSTRACT

The ubiquity of smartphones has made a remarkable influence on everyone's day to day life. Variety of useful built-in sensors provide smartphones with a convenient floor for data collection and analysis. Application development based on the user's location and movement is not a difficult task nowadays. But injuries and deaths due to smartphone-distracted movement on roadways is on the increase. This study explores the capabilities of smartphone inertial sensors for pedestrian activity recognition. Smartphone distracted movements can be predicted from the associated pedestrian's posture, thus inertial sensors can provide effective solution for this specific task. Volunteers were asked to perform different pedestrian activities with smartphones in their hand or in trouser pocket. Accelerometer and gyroscope data were collected, and time windowing was applied for proper segmentation of the data. After time and frequency domain feature extraction of these segmented data streams, two classical supervised machine learning approaches (SVM and Random Forest) were undertaken for correct prediction of seven different pedestrian activity labels. Furthermore, we implemented a deep learning classifier (CNN) for direct activity recognition using raw data. The training and testing procedure includes three types of systems: single-subject, all-subject and leave-one-subject-out models. For performance evaluation, we used the F-score metric, which can reach up to 92.3%, 96.1% and 94.2% for these three models, respectively. CNN with raw data provides much better accuracy than the classical machine learning models. With the capability to identify pedestrian activity and thus distracted pedestrians with great accuracy, our approach lays the foundation for a smartphone application based real time P2V warning system. In this system, the vehicle's driver gets a warning in his smartphone about the nearby presence of a distracted pedestrian.

`

ACKNOWLEDGMENT

I would like to convey my heartiest gratitude to my honorable supervisor Dr. M Abdel-Aty for his excellent supervision and being a constant support in this thesis.

`

TABLE OF CONTENT

`

`

LIST OF FIGURES

`

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Pedestrians, the most vulnerable road users, are 1.5 times more likely than passenger vehicle occupants to be killed in a car crash per trip (USDOT, 2019). In 2019, pedestrian fatalities went up by 5 percent compared to the previous year, with 6590 pedestrians killed. This total interprets as two deaths per hundred thousand people, which is the highest since 1997 (Beck et al., 2007). It is worth noting the fact that smartphone-distraction is responsible for a lot of pedestrian deaths and injuries. According to distracted walking study in 2015, nearly 4 out of 10 Americans say they have personally witnessed a distracted walking incident, and just over a quarter (26%) say they have been in an incident themselves. Pedestrians who text while walking are 60 per cent more likely to veer off-line than non-texters (Safety Team, 2019). Moreover, 60 percent of pedestrians walk while performing smartphone distracted activities like texting, emailing, talking on the phone, or listening to music, despite that 70 percent of the pedestrians consider those behaviors to be dangerous (Insurance, 2013). As pedestrians do not provide any indication to mend their inattentive walking behavior, it is high time we developed tools to mitigate this problem and make roads safer for pedestrians.

The rise in smartphones' distracted walking injuries mirrors the eight-fold increase in cell phone use in the last 15 years. Since pedestrians are getting injured and killed due to these smartphones, why not tap into the sensing power of the smartphone itself for a solution? The development of mobile phone applications based on the user's movement and location is not a difficult task nowadays with smartphone's embedded sensors (Lane et al., 2010), GPS, application programming interfaces (API) and computing ability. We already have large-scale deployment of smartphones in most areas. The percentage of people carrying a smartphone went up from 35% in 2011 to 81% in 2019 (PEW, 2019). Smartphone-distractedness can be modeled by the associated pedestrian's posture. Accelerometer sensor returns acceleration force data for three cartesian coordinate axes, whereas gyroscope provides rate of rotation for three coordinate axes. Since, smartphone built-in accelerometer and gyroscope sensors can pick up pedestrian posture to some extent, they possess the potential to identify distracted pedestrians. Figure 1 shows the triaxial accelerometer and gyroscope raw data collected from smartphones.

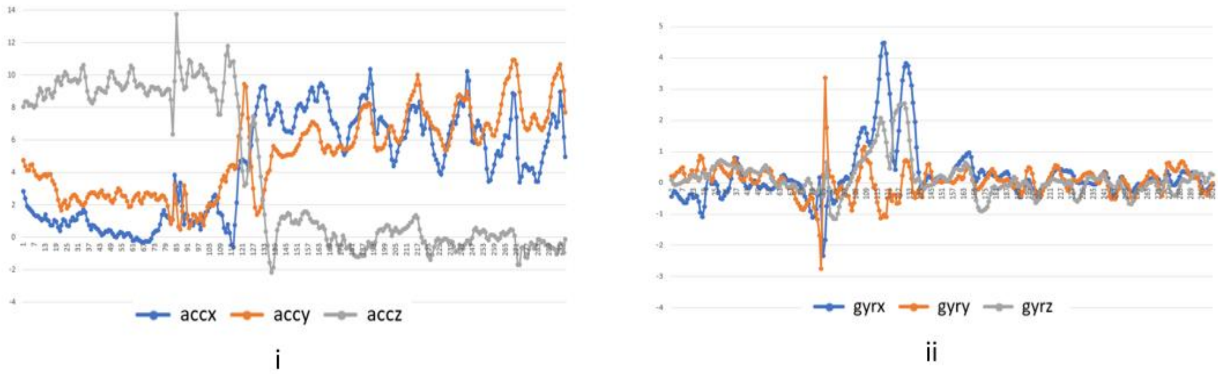


Figure 1 Smartphone inertial triaxial sensor raw data: (i) accelerometer and (ii) gyroscope

This study presents a framework (Figure 2) and evaluation of smartphone-sensor based pedestrian activity recognition, which can detect a pedestrian's distractedness. Accelerometer and gyroscope data were collected when participants perform some typical smartphone-distracted and undistracted pedestrian activities. Then we acquire time and frequency domain features from the data and implement supervised machine learning models to identify pedestrian activity labels. A deep learning model using only raw data has also been implemented. We performed three types of training and testing procedures: single-subject, all-subject and leave-one-subject-out model. Our main purpose of this study is to perform a systematic analysis of sensor-behavior based pedestrian activity recognition via smartphones.

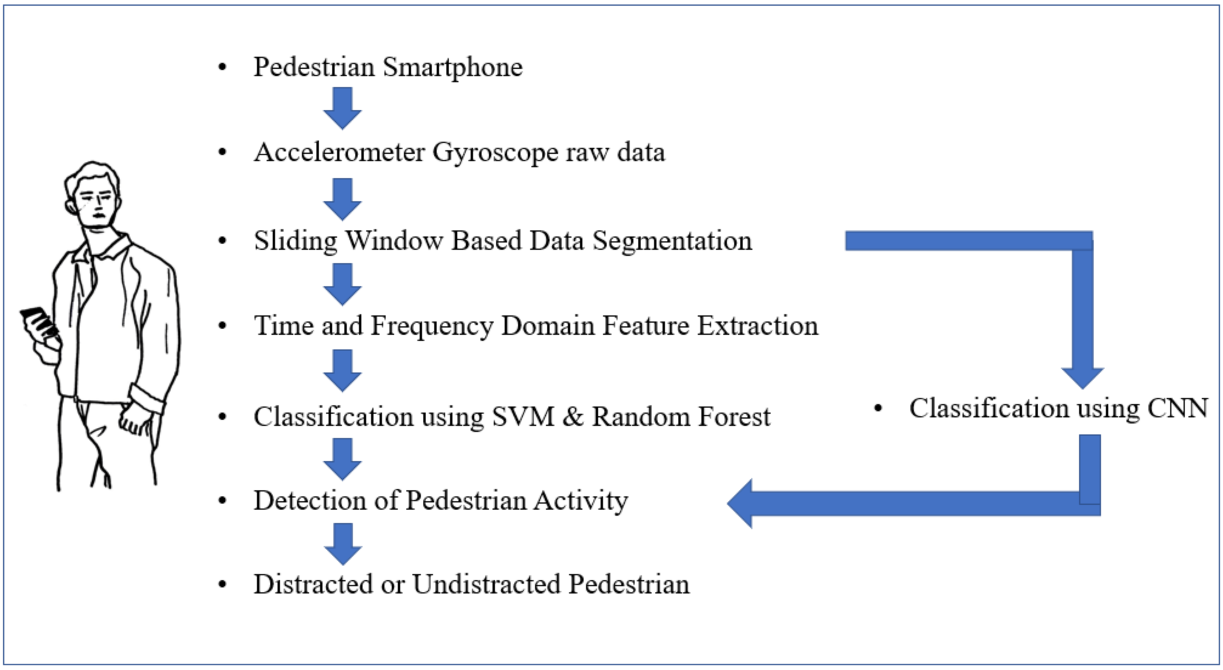


Figure 2 Pedestrian activity recognition system overview

The main contribution of this study is to view smartphone inertial sensors as useful devices to identify pedestrian motions in the context of distractedness. Earlier, smartphone sensors have been utilized to detect daily activities of human life, but this is the first study to recognize smartphone distracted pedestrian behavior using accelerometer and gyroscope sensors for P2V communication and warning systems. In a smart vehicle-pedestrian system, the data collected from pedestrians distracted behavior is sent to the server where the nearby vehicle driver is warned real-time in his smartphone about the presence of a distracted pedestrian nearby.

CHAPTER 2: LITERATURE REVIEW

Smartphone sensors have been used in previous studies for traffic and road condition estimation (Chugh et al., 2014), outdoor position tracking (Zhu et al., 2013), map matching (Bierlaire et al., 2013), estimation of safety performance measures (Guido et al., 2012) and analyzing driving and road events (Kalra et al., 2014).

Wang et al. (2012) have developed a phone camera-based pedestrian safety smartphone app to detect if cars are coming toward the user. Major drawback of this approach is that the phone needs to be held at certain specific positions. CarSafe (You et al. 2013) alerts distracted drivers using dual cameras on smartphones for detecting driver state and for tracking road conditions. The limitation of smartphone camera-based systems is that they fail to provide good output in bad light conditions, and computational load is high for these systems. Smartphone GPS sensors have been proven to be useful in pedestrian risk detection in suburban environments (Jain et al., 2014). Later, their approach has been enhanced by exploring the role of inertial sensors to predict turns made by pedestrians near an intersection (Datta et al., 2015). Chen and Hu (2012) have integrated Inertial Measurement Unit with GPS by designing an Extended Kalman Filter algorithm to predict the position and attitude of a pedestrian. Liu et al. (2016) proposed a driving behavior detection mechanism for commodity smartphones. They broadcast the compressed sensing data by decreasing Wi-Fi association and authentication overhead using the Wi-Fi beacon to notify surrounding drivers. Finally, a collision estimation algorithm is proposed to provide the appropriate warnings.

Smartphone sensing of vehicle dynamics can be utilised to determine driver phone use. Wang et al. (2013) used accelerometers and gyroscopes of the smartphone to detect centripetal acceleration differences on account of vehicle dynamics. When combined with angular speed, these differences can figure out whether the phone is on the left or right side of the vehicle. Wu et al. (2016) used smartphone sensors to detect lane change behavior patterns. Chen et al. (2015) developed a non-vision sensor-based vehicle steering detection middleware on smartphones and proposed algorithms for identification of various vehicle movements and driving on curvy roads considering smooth driving behaviors. Saiprasert et al. (2017) used different smartphone sensors for driving event detection. They proposed one GPS sensor-based detection algorithm and two accelerometer sensor-based pattern matching algorithms.

Sensor-based activity recognition is a well-studied topic. Specialized and smartphone-embedded types of inertial sensors have been implemented in the past. Number of sensors, participants and activities, location of sensors on participant's body, and types of activities can vary among different research groups. Pärkkä et al. (2006) achieved 86% accuracy in classifying seven activities: lying, rowing, riding a bike, standing still, running, walking and Nordic walking. They collected acceleration, vital signs and environmental variable signals from multiple sensors located on the individual's body. After time and frequency domain feature extraction, decision tree and artificial neural network have been applied. Labeling inaccuracy during the transition periods led to most of the misclassifications in their study. Bao & Intille (2004) considered 20 activities and used five bi-axial accelerometers on the user's knee, ankle, arm and hip. Their accuracy was 84% with decision tree classifier using time and frequency domain features. The classifier mainly misclassified some of the complex activities like stretching, scrubbing, riding escalator, and riding elevator. Khan et al. (2010) detected human activities as well as transitions between activities. After manually extracting three sets of features from acceleration data collected by an accelerometer placed on the individual's chest, Linear Discriminant Analysis was applied to reduce feature vector dimension. They found Tilt angle, defined as the angle between the positive Z-axis and the gravitational vector g, to be an important feature for accuracy improvement. Zhu & Sheng (2009) collected acceleration data from a single subject using two accelerometer sensors placed on the subjects wrist and waist. They apply ANN to discriminate among stationary and non-stationary activities, followed by specific activity prediction using Hidden Markov Model (HMM). The system proposed by Randell & Muller (2000) attained 95% accuracy in recognizing ambulation activities. Their method computes Root Mean Square (RMS) from acceleration signals and implements a Backpropagation Neural Network for classification. Vinh et al. (2011) achieved 88.38% accuracy using semi-Markovian conditional random fields to perceive routines that are formulated by successions of subsets of activities from a total collection of 20 activities. These routines include dinner, commuting, lunch and office.

Many researchers have used only smartphone accelerometer data for human activity recognition. Brezmes et al. (2009) implemented a simple yet accurate kNN based six daily activity classifier with cell phone accelerometer data. In their approach, the user is required to train the application on his/her own with the smartphone in any location in his body. Kwapisz et al. (2011) achieved 90 percent accuracy in identifying five daily activities using time domain features. Their collected accelerometer data was from 29 participants and the phone was carried in pant front pockets. Henpraserttae et al. (2011) conducted two experiments with ten subjects. In the first experiment, the phone is fixed on the waist in 16 different orientations, In the second one, the phone is placed in a shirt pocket, trouser-pocket and waist in two different orientations. they utilized mean and standard deviation over triaxial acceleration signals to construct a feature vector and achieved up to 90 percent accuracy with an instance-based classifier (k=3). Sun et al. (2010) implemented an SVM classifier with time and frequency domain features from the acceleration data to get maximum 94.8% overall F-score. They collected data from seven subjects in six pocket positions and used seven activity levels. Wu et al. (2012) found better activity recognition results after incorporation of gyroscope data with acceleration information. They implemented seven classifiers for 16 participants on 13 activities. Helaoui et al. (2011) performed concurrent activity recognition with a statistical-relational approach. Besides these, online Human Activity Recognition (HAR) systems have also been implemented in the past (Berchtold et al., 2010a; Berchtold et al., 2010b; Kao et al., 2009; Lara & Labrador, 2012; Maurer et al., 2006; Riboni & Bettini, 2011).

Recently deep learning-based approaches have been undertaken in studies demonstrating functional accuracy in activity recognition from body-fixed sensor information (Anderson et al., 2007; Győrbíró et al., 2009; Hammerla et al., 2016; Ronao and Cho, 2015; Saponas et al., 2008). In fact, Alsheikh et al. (2016) and Gjoreski et al. (2016) have shown performance improvements with deep learning methods than other state-of-the-art methods.

In this study, we attempt to detect multiple fine-grained simultaneous pedestrian actions like walking and texting, walking and talking, walking with phone in the trouser front pocket etc. with several classical machine learning and deep learning techniques. The concept of posture identification using smartphone inertial sensors is applied to predict the activities.

CHAPTER 3: PRELIMINARY WORK

Before our main work, which aims to detect pedestrian activity based on inertial sensors, we performed a pedestrian trajectory estimation analysis using smartphone GPS points. For data collection, we walked along a path at UCF Gemini Boulevard involving sidewalks and crosswalks for 30 minutes holding a phone in our hand. Then we utilized the data to analyze using Long Short-Term Memory (LSTM) networks and Adaptive filters.

Figure 3 shows a block diagram of the proposed system.

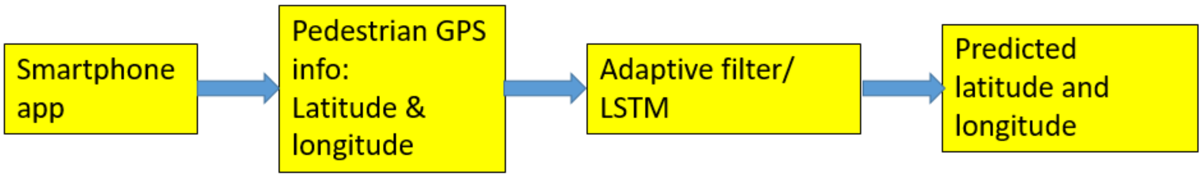


Figure 3 Block diagram of the trajectory prediction system

LSTM networks are a deep and recurrent model of neural networks. Recurrent networks differ from the traditional feed-forward networks in the sense that they do not only have neural connections on a single direction, that means neurons can pass information to a previous or the same layer. So, data does not flow in a single way, and the practical consequences for that is the existence of short-term memory, additional to the long-term memory that neural networks already possess in consequence of training. LSTM were introduced by Hochreiter and Schmidhuber (1997), and it aimed for a superior performance by addressing the vanishing gradient issue that recurrent networks would suffer when dealing with long data sequences.

We use three LSTM models: model 1, 2 and 3 consist of 1, 2 and 3 hidden layers, respectively. The inputs are latitudes and longitudes data, we use Adam optimizer, mean squared error loss function, batch size of 32 and 100 epochs.

An adaptive filter is a filter that self modifies its transfer function according to an optimizing algorithm. It adapts the performance based on the input signal. Such filters incorporate algorithms that allow the filter coefficients to adapt to the signal statics. Least mean squares (LMS) algorithms are class of adaptive filter used to mimic a desired filter by finding the filter coefficients that relate to producing the least mean squares of the error signal (difference between the desired and the actual signal). It is a stochastic gradient descent method in that the filter is only adapted based on the error at the time.

The LSTM model with three hidden layers produced the best results. Table 1 demonstrates the average prediction errors.

Table 1 Performance analysis of different models

Model	Average prediction error
LSTM (model 1)	10.3 feet
LSTM (model 2)	8.4 feet
LSTM (model 3)	7.3 feet
LMS adaptive filter	9.5 feet

Figure 4 shows the real vs predicted trajectories of the best LSTM model and the adaptive filter model.

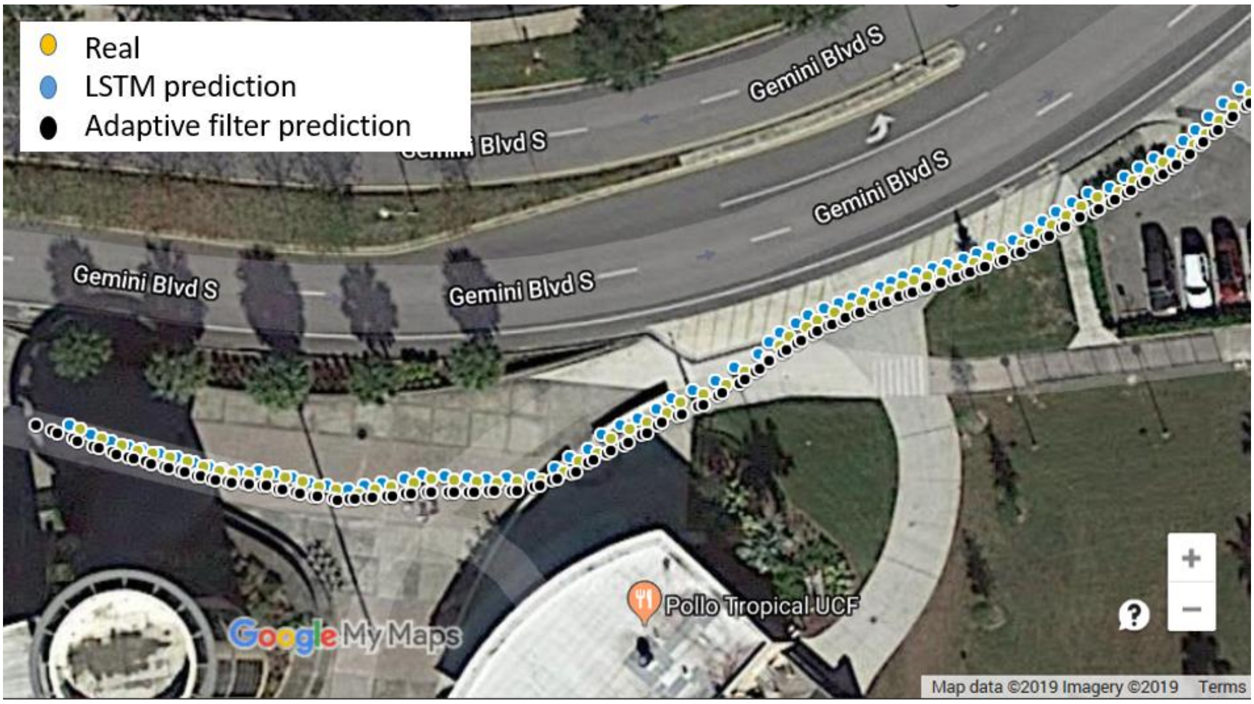


Figure 4 Real vs predicted trajectory (test data)

This study predicts pedestrian trajectory using smartphone localization data. LSTM performed better than adaptive filters in our case. This study implements LSTM and adaptive filter using GPS information without performing any preprocessing of the data at all. Thus, these models are capable of learning directly from latitude and longitude data, not requiring any sort of encoding.

CHAPTER 4: PROBLEM STATEMENT

In this study, we attempt to recognize pedestrian activity from triaxial accelerometer and gyroscope data from smartphones. So, we approach it as a classification problem. In our analysis, we have seven pedestrian activity labels: $\mathbf{P= \{P_1, P_2, P_3, \ldots, P_7\}}$. Figure 2 shows the activity postures considered in this study.

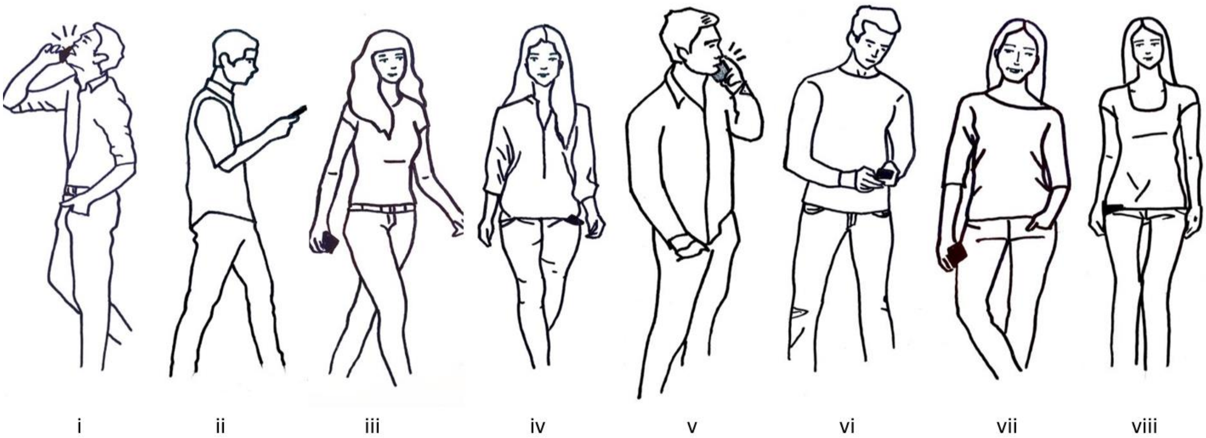


Figure 5 Typical postures for pedestrians: (i) walking and talking, (ii) walking and texting, (iii) walking with phone is hand, (iv) walking with phone in trouser front pocket, (v) standing and talking, (vi) standing and texting, (vii) standing with phone in hand and (viii) standing with phone in trouser front pocket.

The last two postures have been labeled as a single class in our analysis and termed as the ‘standing undistracted’ posture. The series of smartphone sensor readings that captures different pedestrian activity information: $\mathbf{Q= \{Q_1, Q_2, \ldots\ldots\ldots, Q_t, \ldots\ldots, Q_n\}}$. Each Q is a set of nine smartphone sensor signals. We denote triaxial body accelerations, gyroscope and total accelerations by A_x, A_y, A_z, G_x, G_y, G_z, T_x, T_y and T_z. So, $\mathbf{Q_t= \{A_{xt}, A_{yt}, A_{zt}, G_{xt}, G_{yt}, G_{zt}, T_{xt}, T_{yt}}$ **and** $\mathbf{T_{zt}\}}$, subscript t denotes reading at time t.

Now, we need to design a model M to predict the activity sequences based on sensor readings **Q**. If the predicted activity is $\widehat{\boldsymbol{P}}$, then $\widehat{\boldsymbol{P}}\mathbf{=M(Q)}$, $\widehat{\boldsymbol{P}} \in \mathbf{\{P_1, P_2, P_3, \ldots, P_7\}}$. Actual activity, $\mathbf{P^*}$ is the activity ground truth, $P^* \in \mathbf{\{P_1, P_2, P_3, \ldots, P_7\}}$.

Our goal is to learn the model M by minimizing the Cost Function, **C (M(Q), P*).** The positive cost function is defined as the inconsistency between the predicted activity **P'** and then ground truth activity **P*.** To be more specific, We define a function **J** to project the sensor reading data $Q_i \in Q$ to a feature vector **J(Qi)** and then minimize the cost function, **C(M(J(Qi)), P*).**

CHAPTER 5: DATA PREPARATION

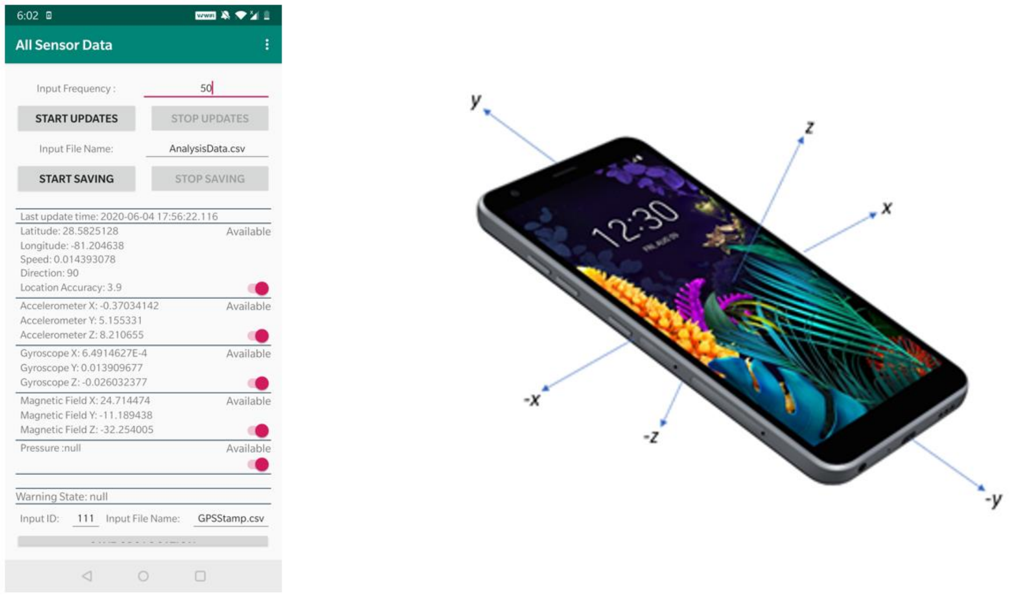
5.1 **Data collection:**

To facilitate the data collection process, we built an android application. The application can continuously display and log triaxial accelerometer, gyroscope and magnetometer data, latitude, longitude and timestamp at our selected frequency (50Hz). The application fetches raw inertial sensor data using Android Sensor Framework. Figure 3 shows a screenshot of the user interface of our application and the Cartesian Coordinate System convention for smartphone inertial sensors. For data collection, we recruited 15 people (age 22-30 years) on the University of Central Florida (UCF) campus, whose heights range from 155cm to 180cm and weights range from 55kg to 80kg (after permission from Institutional Research Board, IRB UCF). Among them, 10 were male and 5 are female. In this study, they used two different smartphones: Samsung Galaxy Note 9 and One Plus 6. Each participant performed eight characteristic of distracted and undistracted pedestrian activities at a sidewalk inside UCF campus: walking and texting, walking and talking over phone, walking and holding the phone in hand (undistracted), walking with phone in trouser pocket, standing and texting, standing and talking over phone, standing and holding the phone in hand, and standing and keeping the phone in pocket (Table 1). The volunteers were asked to perform these actions as naturally as possible and they could take a break whenever they felt like. Some of the them preferred the left hand for holding the phone while performing different activities, whereas others opted for right hand. All of them used portrait mode as the smartphone orientation for the texting scenarios. The collected data were analyzed offline using Python framework.

Table 1: Pedestrian Activity List (each action was performed for about three minutes by each volunteer)

Table 2 Pedestrian Activity List (each action was performed for about three minutes by each volunteer)

Dynamic Activities	**Static activities**
Walking and talking over phone	Standing and talking over phone
Walking and texting	Standing and texting
Walking and holding the phone in hand (undistracted)	Standing and holding the phone in hand (undistracted)
Walking with the phone kept in trouser front pocket	Standing with the phone in trouser front pocket



i

ii

Figure 6(i) Preview of the sensor data collection application and (ii) Cartesian coordinates for smartphone

5.2 **Sampling frequency correction:**

Upon a careful investigation of the data, we found that the sampling frequency to be somewhat inconsistent. The sampling frequency varies between 48-54hz. In order to get a constant frequency, we interpolate and log readings every 0.02 second.

5.3 **Signal processing:**

For noise reduction, we preprocess the acceleration and gyroscope signal with a median filter and a 20 Hz cutoff frequency 3rd order Butterworth low pass filter. This rate is good enough to sense human body motion with reasonable accuracy since 99% of its energy is contained below 15Hz (Karantonis et al., 2006). The acceleration signal contains gravitational and body motion components. The gravitational acceleration possesses low frequency components. In order to separate the gravitational force, we apply a 0.3Hz cutoff frequency low pass filter. So, in our analysis, we have nine signals: total acceleration, body acceleration and body gyroscope values.

5.4 **Windowing:**

Raw sensor data must be divided into segments and tagged to an activity for proper analysis. To quantize the raw data stream, we apply time windowing. Too small window size does not allow robust feature profiling. On the contrary, decision making gets delayed if the window size is too large even though robustness increases in the latter case. Thus, we require an optimum window size.

We also want to make sure one full walking cycle is characterized in each window. Cadence of an individual is between 90 to 130 steps per minute or a minimum of 1.5 steps per sec (BenAbdelkader et al., 2002) . Smartphone-distracted pedestrians tend to walk slower than undistracted pedestrians (Sayed et al., 2013). In this study, we opt for 2.56 sec windows, which means 128 data points per window at 50Hz with 50% overlap among them.

CHAPTER 6: FEATURE SELECTION

For feature selection, we have applied time and frequency domain features: mean, standard deviation, median absolute value, max, min, signal magnitude area, energy, interquartile range, entropy and correlation coefficient (Table 3)

Table 3 Feature Selection

Function	Description
mean	Mean value
std	Standard deviation
mad	Median absolute value
max	Largest value in array
min	Smallest value in array
sma	Signal magnitude area
energy	Average sum of the squares
iqr	Interquartile range
entropy	Signal entropy
correlation	Correlation coefficient

CHAPTER 7: CLASSIFIER IMPLEMENTATION

Considering the nature of our classification problem we pick two classical machine learning models: SVM and random forest and one deep learning classifier: CNN.

7.1 Support vector machine:

SVM attempts to establish the position of decision boundaries which generate the optimal segregation of classes (Cortes and Vapnik, 1995). If the classes are linearly separable in a binary (two-class) classification problem, SVM picks the linear decision boundary that warrants the greatest margin between the two classes. Here, the margin is defined as the sum of the distances to the hyperplane from the closest points or 'support vectors' of the two classes. For the case where the two classes are linearly inseparable, SVM attempts to determine the hyperplane that maximizes the margin and simultaneously minimizes the number of misclassification errors. SVMs can also be expanded to tackle nonlinear decision surfaces. In this case, the input data are projected into a high-dimensional feature space after nonlinear mapping. Thus, a linear classification problem is established in the high space. Kernel functions are applied to decrease the computational complexity of managing high-dimensional feature space.

Even though SVMs were primarily designed for binary problems, they can deal with multi classes with a suitable multi-class method. 'One against one' and 'one against rest' are two frequently used methods for handling multi-class problems. We opt for linear and radial basis function kernels and 'one against rest' approach in this study.

7.2 Random forests:

The Random Forest Ensemble Method uses a combination of tree-type classifiers for classification. These classifier trees are created using a random representative vector from the original input vector, and an input vector is classified according to popular voting by all the trees. Random forests have the flexibility of increasing the number of trees to manage data with high dimensionality. To optimize between performance and required computing power, we select a tree number of 200.

7.3 CNN:

In deep learning, convolutional neural networks are regularized versions of multilayer perceptrons or fully connected networks. CNNs exploit the hierarchical patterns in data and construct more complex patterns with smaller and simpler patterns. The convolutional neural network model used in this study consists of two convolution layers, pooling layer, fully connected layer, dropout layer and SoftMax layer. In convolution layers, depth wise convolution was applied with several convolution filters. Then ReLU activation function was applied to the output of the convolution layer. In the pooling layer max pooling algorithm is used to reduce the spatial size with a filter of size 2×2. Then a fully connected layer is applied followed by dropout layer to avoid overfitting of training data. SoftMax regression was performed in the SoftMax layer at the end. Adam optimizer has been used in this study.

CHAPTER 8: RESULTS

For the evaluation task, we apply three well-known metrices used in pattern recognition: Precision, recall and F_1 score. For each time step, we define the seven activities as the possible labels. True positives (TP) are defined as the correctly predicted activities. A predicted activity p_1 at time step t is considered to be a false positive (FP) if it does not match the activity p_2 in the reference dataset for the same time step. Furthermore, the activity p_2 will be counted as false negative (FN) since it is present in the reference dataset but missing in the prediction for the same time step. Since each time step is tagged with only one activity, the number of false positives and false negatives match.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F_1\ score = \frac{2*Precision*Recall}{Precision+Recall}$$

We go for three different types of training and testing schemes: single-subject, all-subject and leave-one-subject-out analysis.

8.1 Single subject model:

Every pedestrian has unique distinctive movement patterns. Age, weight, height, preference to a specific hand for carrying smartphones, gender, physical condition and weather may have some impact on his/her movement. So, we can analyze each user's data separately. We build single subject personalized models for all the 15 subjects and apply their own data for evaluation. For each volunteer’s data, , we perform four-fold cross-validation by dividing the dataset into four equally sized subsections using only one of them as testing set and the remaining three as training dataset. Finally, we compute the average accuracy. Table 4 shows the results of single-subject models.

Table 4 F1-scores from single-subject model

Participant no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SVM linear	61.45	79.32	71.32	89.32	64.32	78.32	73.32	76.32	91.32	62.32	69.32	71.32	82.32	59.32	81.32
SVM radial	61.78	81.34	71.9	90.23	64.76	76.81	73.34	77.46	90.96	64.34	70.14	71.3	82.1	60.32	82.42
Random Forest	71.94	84.36	72.54	90.91	76.32	81.34	75.35	77.4	92.34	80.14	76.24	75.15	84.32	69.71	85.89

8.2 All-subject model:

Since, every volunteer performs each activity for only three minutes, we do not have the luxury to play with a lot of data in single subject models. In order to increase the training dataset, we merge data from each volunteer. Hence, we train all-subject model using all the 15 volunteer’s data and test only on a single person’s data. After model training, we perform 15 testing experiments for each subject. The F1-scores are shown in Table 5. Table 4-7 demonstrate the confusion matrices for different classifers.

Table 5 F1-scores from All-subject model

Participant no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SVM linear	84.3	89.1	78.2	69.9	81.6	85.2	84.6	78.5	82.7	88.3	71.3	79.5	90.2	79.8	82.1
SVM radial	85.6	86.1	78.7	69.8	82.1	86.2	87.9	79.1	83.9	88.6	73.1	78.8	91.3	79.8	84.7
Random forest	86.6	89.3	79.7	74.8	85.1	88.2	87.8	85.1	84.1	90.3	78.4	79.2	94.3	81.6	89.7
CNN with raw data	88.6	89.1	90.7	92.8	94.1	98.2	90.9	92.1	91.9	92.3	94.1	98.2	96.3	91.8	92.7

Table 6 Confusion Matrix for SVM with linear kernel in all-subject model

		Predicted Activity							
		WT	WTx	WH	WP	ST	STx	SU	Recall
ACTUAL ACTIVITY	Walking and Talking	84	3	14	1	0	0	0	82%
	Walking and Texting	7	93	8	10	0	0	0	79%
	Walking and Holding phone	5	6	105	15	0	0	0	80%
	Walking with phone in Pocket	2	4	10	82	0	0	0	83%
	Standing and Talking	0	0	0	1	127	1	15	88%
	Standing and Texting	0	0	0	0	4	126	18	85%
	Standing Undistracted	0	0	0	0	4	20	224	90%
	Precision	86%	88%	77%	75%	94%	85%	87%	

Table 7 Confusion Matrix for SVM with RBF kernel in all-subject model

		Predicted Activity							
		WT	WTx	WH	WP	ST	STx	SU	Recall
ACTUAL ACTIVITY	Walking and Talking	91	5	6	0	0	0	0	89%
	Walking and Texting	7	97	5	9	0	0	0	82%
	Walking and Holding phone	5	6	112	8	0	0	0	83%
	Walking with phone in Pocket	2	6	8	82	0	0	0	84%
	Standing and Talking	0	0	0	0	134	1	9	93%
	Standing and Texting	0	0	0	0	1	132	15	89%
	Standing Undistracted	0	0	0	0	2	18	228	92%
	Precision	86%	85%	85%	83%	98%	87%	90%	

Table 8 Confusion Matrix for Random Forest in all-subject model

		Predicted Activity							
		WT	WTx	WH	WP	ST	STx	SU	Recall
ACTUAL ACTIVITY	Walking and Talking	90	2	9	1	0	0	0	88%
	Walking and Texting	5	95	7	11	0	0	0	81%
	Walking and Holding phone	5	6	109	11	0	0	0	83%
	Walking with phone in Pocket	2	4	10	82	0	0	0	84%
	Standing and Talking	0	0	0	0	131	0	13	91%
	Standing and Texting	0	0	0	0	1	129	18	87%
	Standing Undistracted	0	0	0	0	4	15	229	92%
	Precision	88%	89%	81%	78%	96%	89%	88%	

Table 9 Confusion Matrix for CNN in all-subject model

		Predicted Activity							
		WT	WTx	WH	WP	ST	STx	SU	Recall
ACTUAL ACTIVITY	Walking and Talking	97	4	1	0	0	0	0	95%
	Walking and Texting	1	116	0	1	0	0	0	98%
	Walking and Holding phone	1	2	128	0	0	0	0	98%
	Walking with phone in Pocket	1	0	3	94	0	0	0	96%
	Standing and Talking	0	0	0	0	135	0	9	94%
	Standing and Texting	0	0	0	0	1	135	12	91%
	Standing Undistracted	0	0	0	0	4	7	237	95%
	Precision	97%	95%	97%	99%	96%	95%	92%	

8.3 Leave-one-subject-out model:

This is like the all-subject model, but the training data omits the tested subject's data. The purpose for this type of analysis is the fact that we are interested in identifying if a trained model can predict a new subject's activity with reasonable accuracy. If the answer is affirmative, then already trained classifiers can be used to identify an unknown pedestrian's activity. The F1-scores for this study are demonstrated in Table 10 and table 8-11 show the confusion matrices for different classifiers

Table 10 F1-scores for Leave-one-subject-out model

Participant no.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SVM linear	78.3	68.1	71.2	68.9	74.6	79.2	65.6	79.5	80.7	79.3	70.3	71.5	86.2	72.8	79.1
SVM radial	79.3	68.2	71.9	69.1	75	79.6	66.4	79.7	81.7	79.2	70.9	72	86.9	72.9	80
Random forest	84.6	81.3	78.7	73.8	83.1	88.2	87.8	83.1	84.1	89.3	74.4	78.2	90.3	78.6	86.7
CNN with raw data	86.6	85.1	90.6	91.7	90.1	95.2	89.8	91	90.4	87.3	93.6	98.3	95.1	90.6	87.3

Table 11 Confusion Matrix for SVM with linear kernel in leave-one-subject-out model

		Predicted Activity							
		WT	WTx	WH	WP	ST	STx	SU	Recall
ACTUAL ACTIVITY	Walking and Talking	81	6	14	1	0	0	0	79%
	Walking and Texting	7	92	8	11	0	0	0	85%
	Walking and Holding phone	5	10	101	15	0	0	0	77%
	Walking with phone in Pocket	2	4	11	81	0	0	0	83%
	Standing and Talking	0	0	0	1	121	1	21	84%
	Standing and Texting	0	0	0	0	4	125	19	85%
	Standing Undistracted	0	0	0	0	4	21	223	90%
	Precision	85%	82%	75%	78%	94%	85%	84%	

Table 12 Confusion Matrix for SVM with RBF kernel in leave-one-subject-out model

		Predicted Activity							
		WT	WTx	WH	WP	ST	STx	SU	Recall
ACTUAL ACTIVITY	Walking and Talking	84	5	6	7	0	0	0	82%
	Walking and Texting	11	93	5	9	0	0	0	79%
	Walking and Holding phone	4	6	103	18	0	0	0	79%
	Walking with phone in Pocket	2	6	8	82	0	0	0	84%
	Standing and Talking	0	0	0	2	120	1	21	83%
	Standing and Texting	0	0	0	0	1	120	27	81%
	Standing Undistracted	0	0	0	0	2	25	220	89%
	Precision	83%	85%	84%	70%	97%	82%	82%	

Table 13 Confusion Matrix for Random Forest in leave-one-subject-out model

		Predicted Activity							
		WT	WTx	WH	WP	ST	STx	SU	Recall
ACTUAL ACTIVITY	Walking and Talking	86	6	9	1	0	0	0	84%
	Walking and Texting	5	94	8	11	0	0	0	80%
	Walking and Holding phone	5	6	105	15	0	0	0	80%
	Walking with phone in Pocket	2	4	10	82	0	0	0	84%
	Standing and Talking	0	0	0	0	130	1	13	90%
	Standing and Texting	0	0	0	0	4	126	18	85%
	Standing Undistracted	0	0	0	0	4	16	228	92%
	Precision	88%	81%	79%	75%	94%	85%	88%	

Table 14 Confusion Matrix for CNN in leave-one-subject-out model

		Predicted Activity							
		WT	WTx	WH	WP	ST	STx	SU	Recall
ACTUAL ACTIVITY	Walking and Talking	93	7	2	0	0	0	0	91%
	Walking and Texting	2	105	10	1	0	0	0	89%
	Walking and Holding phone	7	3	116	5	0	0	0	89%
	Walking with phone in Pocket	1	0	6	91	0	0	0	93%
	Standing and Talking	0	0	0	0	127	13	4	88%
	Standing and Texting	0	0	0	0	10	136	2	92%
	Standing Undistracted	0	0	0	0	5	9	234	94%
	Precision	90%	86%	86%	95%	88%	92%	95%	

8.4 Discussions:

All subject models and leave one subject out models are more stable than the single subject models. Single subject models lack accuracy due to the inadequacy of data. Having more data points is more important than personality traits. An adequate amount of personalized data should be gathered. CNN using raw data provides best results for both all-subject and leave-one-subject-out models.

Confusion matrices show that models easily separate dynamic and static activities, but they often get confused among different sub-categories within dynamic and static activities.

All-subject model provides the best results. Personalized data are required to promote the recognition accuracy, while generalized data stabilizes the model against unusual training data. Moreover, training data from other users makes the model more robust.

CHAPTER 9: P2V WARNING SYSTEM FOR IDENTIFYING DISTRACTED PEDESTRIANS

The main implementation of this work would be in the form of a P2V smartphone application. The vehicle driver would be warned in a smartphone app if a distracted pedestrian is walking or standing nearby. For this, GPS sensor needs to be involved besides the inertial sensors. If a pedestrian's activity is detected as distracted within a specific radius around a car, the driver will get a warning in his/her smartphone about a potentially dangerous situation.

We have already done some preliminary tests using smartphone as an OBU emulator in a single vehicle and single pedestrian application and implemented a real time smartphone application-based warning system. If there is a potential conflict between the forthcoming movements of the pedestrian and the vehicle, they both get a warning in their smartphone application. They system considers the lane changing of a vehicle with the smartphone located inside the vehicle. It works like an on-board unit.

We have analyzed the battery consumption and latency of the system. For Samsung note 9, battery consumptions were 255mAh per hour at a sampling frequency of 2Hz with all the sensors activated. GPS sensor consumes the highest battery energy.

The overall delay in any communication technology is a resultant of four components: transmission delay, propagation delay, queuing delay, and processing delay (Ashraf et al., 2017; Martin et al., 1997). The uploading duration is the time difference between the starting time to upload data from the app and the receiving time of data in the server. The downloading duration is the time difference between the data generating time at the server and the receiving time at the app. From our experiments, the average uploading latencies are between 28ms and 33ms while the average downloading latencies are between 75ms and 130ms for Samsung Note 9. Latency requirement for V2X (vehicle-to-everything) for safety is between 100ms and 1 second (Dey et al., 2016; Xu et al., 2017). Therefore, the average 4G LTE communication latency range of the developed application can meet the latency requirement. GPS accuracy measurement experiments have also provided satisfying results for both pedestrians and vehicles.

CHAPTER 10: CONCLUSIONS

The pervasiveness of smartphones along with its powerful embedded sensors and rapidly improving computational power makes it the perfect platform for tasks like activity recognition. This study presents a machine learning and deep learning-based framework using smartphone inertial sensory data for identification of different pedestrian actions in the context of distractedness. The dataset used in this study contains samples from 15 volunteers, and time and frequency-domain features were extracted to accurately stamp a subject's activity pattern. To perform an organized evaluation, random forest and SVM multiclass classifiers have been implemented after handpicked feature extraction from raw accelerometer and gyroscope data. CNN using raw data have also been implemented. Three different types of training and testing processes have been executed and evaluated. We used handpicked features for this study. In the future, we can consider more features and perform feature selection algorithms like principle component analysis. We can also explore the magnetometer sensor's efficacy in pedestrian activity recognition.

REFERENCES

Alsheikh, M.A., Selim, A., Niyato, D., Doyle, L., Lin, S., Tan, H.P., 2016. Deep activity recognition models with triaxial accelerometers, in: AAAI Workshop - Technical Report.

Anderson, I., Maitland, J., Sherwood, S., Barkhuus, L., Chalmers, M., Hall, M., Brown, B., Muller, H., 2007. Shakra: Tracking and sharing daily activity levels with unaugmented mobile phones. Mob. Networks Appl. https://doi.org/10.1007/s11036-007-0011-7

Ashraf, M.I., Liu, C.F., Bennis, M., Saad, W., 2017. Towards low-latency and ultra-reliable vehicle-to-vehicle communication, in: EuCNC 2017 - European Conference on Networks and Communications. https://doi.org/10.1109/EuCNC.2017.7980743

Bao, L., Intille, S.S., 2004. Activity recognition from user-annotated acceleration data. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). https://doi.org/10.1007/978-3-540-24646-6_1

Beck, L.F., Dellinger, A.M., O'Neil, M.E., 2007. Motor vehicle crash injury rates by mode of travel, United States: Using exposure-based methods to quantify differences. Am. J. Epidemiol. https://doi.org/10.1093/aje/kwm064

BenAbdelkader, C., Cutler, R., Davis, L., 2002. Stride and cadence as a biometric in automatic person identification and verification, in: Proceedings - 5th IEEE International Conference on Automatic Face Gesture Recognition, FGR 2002. https://doi.org/10.1109/AFGR.2002.1004182

Berchtold, M., Budde, M., Gordon, D., Schmidtke, H.R., Beigl, M., 2010a. ActiServ: Activity recognition service for mobile phones, in: Proceedings - International Symposium on Wearable Computers, ISWC. https://doi.org/10.1109/ISWC.2010.5665868

Berchtold, M., Budde, M., Schmidtke, H.R., Beigl, M., 2010b. An extensible modular recognition concept that makes activity recognition practical, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and

Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-642-16111-7_46

Bierlaire, M., Chen, J., Newman, J., 2013. A probabilistic map matching method for smartphone GPS data. Transp. Res. Part C Emerg. Technol. https://doi.org/10.1016/j.trc.2012.08.001

Brezmes, T., Gorricho, J.L., Cotrina, J., 2009. Activity recognition from accelerometer data on a mobile phone, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-642-02481-8_120

Chen, D., Cho, K.T., Han, S., Jin, Z., Shin, K.G., 2015. Invisible sensing of vehicle steering with smartphones, in: MobiSys 2015 - Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services. https://doi.org/10.1145/2742647.2742659

Chen, L., Hu, H., 2012. IMU/GPS based pedestrian localization, in: 2012 4th Computer Science and Electronic Engineering Conference, CEEC 2012 - Conference Proceedings. https://doi.org/10.1109/CEEC.2012.6375373

Chugh, G., Bansal, D., Sofat, S., 2014. Road Condition Detection Using Smartphone Sensors: A Survey. Int. J. Electron. Electr. Eng.

Cortes, C., Vapnik, V., 1995. Support-Vector Networks. Mach. Learn. https://doi.org/10.1023/A:1022627411411

Datta, T., Jain, S., Gruteser, M., 2015. Towards city-scale smartphone sensing of potentially unsafe pedestrian movements, in: Proceedings - 11th IEEE International Conference on Mobile Ad Hoc and Sensor Systems, MASS 2014. https://doi.org/10.1109/MASS.2014.106

Dey, K.C., Rayamajhi, A., Chowdhury, M., Bhavsar, P., Martin, J., 2016. Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication in a heterogeneous wireless network - Performance evaluation. Transp. Res. Part C Emerg. Technol. https://doi.org/10.1016/j.trc.2016.03.008

Gjoreski, H., Bizjak, J., Gjoreski, M., Gams, M., 2016. Comparing Deep and Classical Machine Learning Methods for Human Activity Recognition using Wrist Accelerometer. 25Th Int. Jt.

Guido, G., Vitale, A., Astarita, V., Saccomanno, F., Giofré, V.P., Gallelli, V., 2012. Estimation of Safety Performance Measures from Smartphone Sensors. Procedia - Soc. Behav. Sci. https://doi.org/10.1016/j.sbspro.2012.09.824

Győrbíró, N., Fábián, Á., Hományi, G., 2009. An activity recognition system for mobile phones. Mob. Networks Appl. https://doi.org/10.1007/s11036-008-0112-y

Hammerla, N.Y., Halloran, S., Plötz, T., 2016. Deep, convolutional, and recurrent models for human activity recognition using wearables, in: IJCAI International Joint Conference on Artificial Intelligence.

Helaoui, R., Niepert, M., Stuckenschmidt, H., 2011. Recognizing interleaved and concurrent activities: A statistical-relational approach, in: 2011 IEEE International Conference on Pervasive Computing and Communications, PerCom 2011. https://doi.org/10.1109/PERCOM.2011.5767586

Henpraserttae, A., Thiemjarus, S., Marukatat, S., 2011. Accurate activity recognition using a mobile phone regardless of device orientation and location, in: Proceedings - 2011 International Conference on Body Sensor Networks, BSN 2011. https://doi.org/10.1109/BSN.2011.8

Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Comput. https://doi.org/10.1162/neco.1997.9.8.1735

Insurance, liberty M., 2013. New Study Shows Three out of Five Pedestrians Prioritize Smartphones over Safety When Crossing Streets [WWW Document]. URL https://www.libertymutualgroup.com/about-lm/news/articles/new-study-shows-three-out-five-pedestrians-prioritize-smartphones-over-safety-when-crossing-streets

Jain, S., Borgiattino, C., Ren, Y., Gruteser, M., Chen, Y., 2014. On the limits of positioning-based pedestrian risk awareness, in: MARS 2014 - Proceedings of the 2014 Workshop for Mobile Augmented Reality and Robotic Technology-Based Systems, Co-Located with MobiSys 2014. https://doi.org/10.1145/2609829.2609834

Kalra, N., Chugh, G., Bansal, D., 2014. Analyzing Driving and Road Events via Smartphone. Int. J. Comput. Appl. https://doi.org/10.5120/17233-7561

Kao, T.P., Lin, C.W., Wang, J.S., 2009. Development of a portable activity detector for daily activity recognition, in: IEEE International Symposium on Industrial Electronics. https://doi.org/10.1109/ISIE.2009.5222001

Karantonis, D.M., Narayanan, M.R., Mathie, M., Lovell, N.H., Celler, B.G., 2006. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Trans. Inf. Technol. Biomed. https://doi.org/10.1109/TITB.2005.856864

Khan, A.M., Lee, Y.K., Lee, S.Y., Kim, T.S., 2010. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. IEEE Trans. Inf. Technol. Biomed. https://doi.org/10.1109/TITB.2010.2051955

Kwapisz, J.R., Weiss, G.M., Moore, S.A., 2011. Activity recognition using cell phone accelerometers. ACM SIGKDD Explor. Newsl. https://doi.org/10.1145/1964897.1964918

Lane, N.D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., Campbell, A.T., 2010. A survey of mobile phone sensing. IEEE Commun. Mag. https://doi.org/10.1109/MCOM.2010.5560598

Lara, Ó.D., Labrador, M.A., 2012. A mobile platform for real-time human activity recognition, in: 2012 IEEE Consumer Communications and Networking Conference, CCNC'2012. https://doi.org/10.1109/CCNC.2012.6181018

Liu, Z., Wu, M., Zhu, K., Zhang, L., 2016. SenSafe: A smartphone-based traffic safety framework by sensing vehicle and pedestrian behaviors. Mob. Inf. Syst. https://doi.org/10.1155/2016/7967249

Martin, R.P., Vahdat, A.M., Culler, D.E., Anderson, T.E., 1997. Effects of communication latency, overhead, and bandwidth in a cluster architecture, in: Conference Proceedings - Annual International Symposium on Computer Architecture, ISCA. https://doi.org/10.1145/384286.264146

Maurer, U., Smailagic, A., Siewiorek, D.P., Deisher, M., 2006. Activity recognition and monitoring using multiple sensors on different body positions, in: Proceedings - BSN 2006: International Workshop on Wearable and Implantable Body Sensor Networks. https://doi.org/10.1109/BSN.2006.6

Pärkkä, J., Ermes, M., Korpipää, P., Mäntyjärvi, J., Peltola, J., Korhonen, I., 2006. Activity classification using realistic data from wearable sensors. IEEE Trans. Inf. Technol. Biomed. https://doi.org/10.1109/TITB.2005.856863

PEW, 2019. Mobile Fact Sheet [WWW Document]. URL https://www.pewresearch.org/internet/fact-sheet/mobile/

Randell, C., Muller, H., 2000. Context awareness by analysing accelerometer data. Int. Symp. Wearable Comput. Dig. Pap. https://doi.org/10.1109/iswc.2000.888488

Riboni, D., Bettini, C., 2011. COSAR: Hybrid reasoning for context-Aware activity recognition. Pers. Ubiquitous Comput. https://doi.org/10.1007/s00779-010-0331-7

Ronao, C.A., Cho, S.B., 2015. Deep convolutional neural networks for human activity recognition with smartphone sensors, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).

https://doi.org/10.1007/978-3-319-26561-2_6

Safety Team, 2019. Distracted Walking a Major Pedestrian Safety Concern [WWW Document]. URL https://www.safety.com/distracted-walking-a-major-pedestrian-safety-concern/

Saiprasert, C., Pholprasit, T., Thajchayapong, S., 2017. Detection of Driving Events using Sensory Data on Smartphone. Int. J. Intell. Transp. Syst. Res. https://doi.org/10.1007/s13177-015-0116-5

Saponas, T., Lester, J., Froehlich, J., Fogarty, J., Landay, J., 2008. iLearn on the iPhone : Real-Time Human Activity Classification on Commodity Mobile Phones, University of Washington CSE Tech Report UWCSE080402.

Sayed, T., Zaki, M.H., Autey, J., 2013. Automated safety diagnosis of vehicle-bicycle interactions using computer vision analysis. Saf. Sci. https://doi.org/10.1016/j.ssci.2013.05.009

Sun, L., Zhang, D., Li, B., Guo, B., Li, S., 2010. Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-642-16355-5_42

USDOT, 2019. Traffic Safety Facts 2017 Data: Pedestrians. Washington, DC.

Vinh, L.T., Lee, S., Le, H.X., Ngo, H.Q., Kim, H. Il, Han, M., Lee, Y.K., 2011. Semi-Markov conditional random fields for accelerometer-based activity recognition. Appl. Intell. https://doi.org/10.1007/s10489-010-0216-5

Wang, T., Cardone, G., Corradi, A., Torresani, L., Campbell, A.T., 2012. WalkSafe: A pedestrian safety app for mobile phone users who walk and talk while crossing roads, in: HotMobile 2012 - 13th Workshop on Mobile Computing Systems and Applications. https://doi.org/10.1145/2162081.2162089

Wang, Y., Yang, J., Liu, H., Chen, Y., Gruteser, M., Martin, R.P., 2013. Sensing vehicle dynamics for determining driver phone use, in: MobiSys 2013 - Proceedings of the 11th Annual International Conference on Mobile Systems, Applications, and Services. https://doi.org/10.1145/2462456.2464447

Wu, W., Dasgupta, S., Ramirez, E.E., Peterson, C., Norman, G.J., 2012. Classification accuracies of physical activities using smartphone motion sensors. J. Med. Internet Res. https://doi.org/10.2196/jmir.2208

Wu, Z., Li, J., Yu, J., Zhu, Y., Xue, G., Li, M., 2016. L3: Sensing driving conditions for vehicle lane-level localization on highways, in: Proceedings - IEEE INFOCOM. https://doi.org/10.1109/INFOCOM.2016.7524436

Xu, Z., Li, X., Zhao, X., Zhang, M.H., Wang, Z., 2017. DSRC versus 4G-LTE for connected vehicle applications: A study on field experiments of vehicular communication performance. J. Adv. Transp. https://doi.org/10.1155/2017/2750452

You, C.W., Lane, N.D., Chen, F., Wang, R., Chen, Z., Bao, T.J., Montes-de-Oca, M., Cheng, Y., Lint, M., Torresani, L., Campbell, A.T., 2013. CarSafe App: Alerting drowsy and distracted drivers using dual cameras on smartphones, in: MobiSys 2013 - Proceedings of the 11th Annual International Conference on Mobile Systems, Applications, and Services. https://doi.org/10.1145/2462456.2465428

Zhu, C., Sheng, W., 2009. Human daily activity recognition in robot-assisted living using multi-sensor fusion, in: Proceedings - IEEE International Conference on Robotics and Automation. https://doi.org/10.1109/ROBOT.2009.5152756

Zhu, X., Li, Q., Chen, G., 2013. APT: Accurate outdoor pedestrian tracking with smartphones, in: Proceedings - IEEE INFOCOM. https://doi.org/10.1109/INFCOM.2013.6567057